

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2005 年 7 月 21 日 (21.07.2005)

PCT

(10) 国際公開番号
WO 2005/066796 A1

(51) 国際特許分類⁷: G06F 12/08, 12/12
(21) 国際出願番号: PCT/JP2004/019102
(22) 国際出願日: 2004 年 12 月 21 日 (21.12.2004)
(25) 国際出願の言語: 日本語
(26) 国際公開の言語: 日本語
(30) 優先権データ:
特願 2003-424042
2003 年 12 月 22 日 (22.12.2003) JP

(71) 出願人 (米国を除く全ての指定国について): 松下電
器産業株式会社 (MATSUSHITA ELECTRIC INDUS-
TRIAL CO., LTD.) [JP/JP]; 〒5718501 大阪府門真市大
字門真 1 0 0 6 番地 Osaka (JP).

(72) 発明者; および
(75) 発明者/出願人 (米国についてのみ): 中西 龍太
(NAKANISHI, Ryuta). 岡林 はづき (OKABAYASHI,
Hazuki). 田中 哲也 (TANAKA, Tetsuya). 宮阪 修二
(MIYASAKA, Shuji).

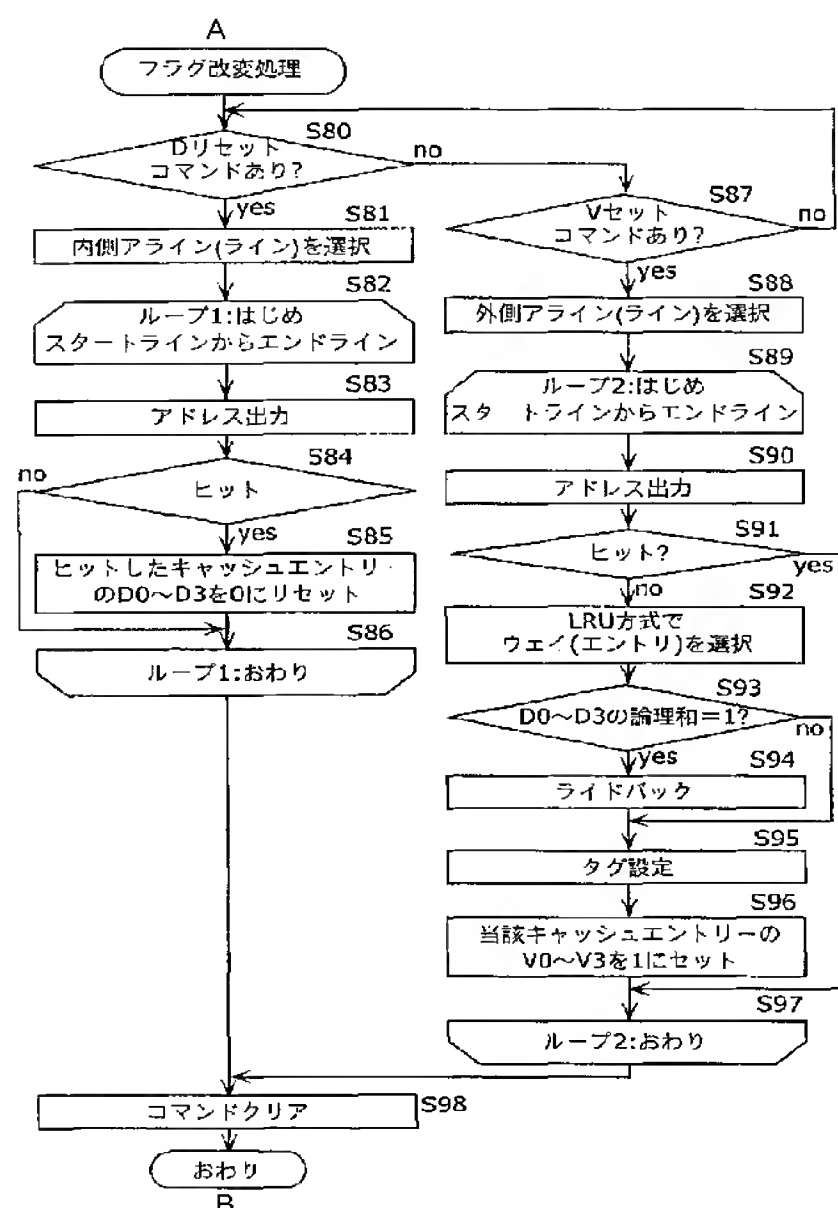
(74) 代理人: 新居 広守 (NII, Hiromori); 〒5320011 大阪府
大阪市淀川区西中島 3 丁目 1 1 番 2 6 号 新大阪末広
センタービル 3 F 新居国際特許事務所内 Osaka (JP).

(81) 指定国 (表示のない限り、全ての種類の国内保護が
可能): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR,
BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM,
DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU,
ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA,
NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE,

[続葉有]

(54) Title: CACHE MEMORY AND ITS CONTROLLING METHOD

(54) 発明の名称: キャッシュメモリ及びその制御方法



A... FLAG ALTERING PROCESSING
S80... D RESET COMMAND?
S81... SELECT INSIDE ALIGN (LINE)
S82... START LOOP 1
FROM START LINE TO END LINE
S83... OUTPUT ADDRESS
S84... HIT
S85... RESET D0 TO D3 OF HIT CACHE ENTRY AT 0
S86... END LOOP 1
S87... V SET COMMAND?
S88... SELECT OUTSIDE ALIGN (LINE)

S89... START LOOP 2
FROM START LINE TO END LINE
S90... OUTPUT ADDRESS
S91... HIT?
S92... SELECT WAY (ENTRY) BY LRU METHOD
S93... LOGICAL SUM OF D0 TO D3=1?
S94... WRITE BACK
S95... SET TAG
S96... SET V0 TO V3 OF THE CACHE ENTRY AT 1
S97... END LOOP 2
S98... CLEAR COMMAND
B... END

(57) Abstract: A cache memory has a valid flag indicating whether a cache entry holding unit data on cache is valid or not correspondingly to the cache entry and a dirty flag indicating whether write for the cache entry is done or not. The cache memory comprises a flag altering section for setting a valid flag by creating the address of the cache entry as a tag without loading data from a memory according to the instruction from the processor or resetting the dirty flag while holding rewritten data not written back for the cache entry.

(57) 要約: 本発明のキャッシュメモリは、キャッシュの単位データを保持するキャッシュエントリーに対応させて、当該キャッシュエントリーが有効であるか否かを示すバリッドフラグと、当該キャッシュエントリーに対する書き込みがなされたか否かを示すダーティフラグと有するキャッシュメモリであって、プロセッサからの指示に基づき、メモリからデータをロードすることなく、キャッシュエントリーに対してタグとしてのアドレスを設定しバリッドフラグをセットし、あるいは、キャッシュエントリーに対してライトバックされていない書き換えられたデータを保持している状態でダーティフラグをリセットするフラグ改変部を備える。

WO 2005/066796 A1



SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, YU, ZA, ZM, ZW.

OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML,
MR, NE, SN, TD, TG).

(84) 指定国 (表示のない限り、全ての種類の広域保護
が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA,
SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ,
BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE,
BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU,
IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR),

添付公開書類:

— 国際調査報告書

2文字コード及び他の略語については、定期発行される
各PCTガゼットの巻頭に掲載されている「コードと略語
のガイダンスノート」を参照。

明 細 書

キャッシュメモリ及びその制御方法

技術分野

- [0001] 本発明は、プロセッサのメモリアクセスを高速化するためのキャッシュメモリおよびその制御方法に関する。

背景技術

- [0002] 近年のマイクロプロセッサでは、例えば、SRAM (Static Random Access Memory) 等から成る小容量で高速なキャッシュメモリをマイクロプロセッサの内部、もしくはその近傍に配置し、データの一部をキャッシュメモリに記憶することによって、マイクロプロセッサのメモリアクセスを高速化させている。
- [0003] コンピュータシステムでは、中央処理装置からキャッシュメモリへのリードアクセスまたはライトアクセスがミスヒットした場合に、主記憶装置から新たに読み出されたデータの一部が、エントリ(登録項目)としてキャッシュメモリの空きブロックに格納される。この時、空きブロックが存在しない場合には、複数のブロックのいずれか1つを選択し、選択されたブロックに格納されているエントリを主記憶装置に戻してブロック内を空き状態にし、この空きブロックに新たに読み出したデータを格納するエントリ置換処理が必要になる。上記エントリ置換処理では、最も以前に参照したデータを格納しているブロックを選択する手法、すなわち、LRU (Least Recently Used) デコード方式が一般的に採用されている。このLRUデコード方式によってキャッシュメモリの使用効率が向上し、その結果、マイクロプロセッサの実行速度が向上する。
- [0004] マイクロプロセッサが処理するプログラムの中には、アクセス頻度は少ないものの、ひとたび起動された場合には高速に処理しなければならないような特殊な処理と、アクセス頻度は多いが、実行速度がそれほど要求されないような処理とが存在する。
- [0005] そこで、これに対応するために例えば特許文献1等の従来技術では、キャッシュメモリにフリーズ機能を設けている。フリーズ機能は、アクセスは少ないものの、ひとたび起動された場合には高速に処理しなければならないようなプログラムを予めキャッシュメモリ内にコピーしておき、その領域を書き換え禁止にしておく機能である。この

機能を有することで、コンピュータシステムは、必要な時にプログラムをキャッシュメモリから読み出して実行することができ、これにより実行時間が短縮する。また、パージ機能は、アクセス頻度は多いが、実行速度がそれほど要求されないようなプログラムやデータをキャッシュメモリ内に保存しておくことなく、その領域を解放する機能である。この機能を有することで、キャッシュメモリに余裕ができ、優先度の高い他のプログラムやデータをキャッシュメモリに取り込むことができ、これにより、キャッシュメモリの利用効率が向上し、総合的な実行時間が短縮する。

特許文献1:特開2000-200221号公報

発明の開示

発明が解決しようとする課題

- [0006] しかしながら、フリーズ機能を使用してもパージ機能を使用しても、キャッシュミス発生によりリプレースする場合に無駄なリプレースをしたり、無駄なライトバックをする場合があるという問題がある。
- [0007] 例えば、無駄なリプレースが生じる場合として、プロセッサが配列要素を全てライトする場合がある。この場合、メモリから新たにデータをキャッシュメモリにリプレースしても、プロセッサから全てライトされるので、リプレースが無駄に終わる。
- [0008] また、無駄なライトバックが生じる場合として、キャッシュエントリーが単にワークデータとして使用される場合がある。この場合、最終的に破棄してもよいデータであるにも拘らず、無駄なライトバックが発生する。

課題を解決するための手段

- [0009] 本発明の目的は、無駄なリプレースや無駄なライトバックを防止するキャッシュメモリを提供することにある。
- [0010] 上記課題を解決するため本発明のキャッシュメモリは、キャッシュの単位データを保持するキャッシュエントリーに対応させて、当該キャッシュエントリーが有効であるか否かを示すバリッドフラグと、当該キャッシュエントリーに対する書き込みがなされたか否かを示すダーティフラグとを保持するフラグ保持手段と、プロセッサから指示に基づき、前記バリッドフラグおよびダーティフラグの少なくとも一方をキャッシュエントリーの状態に反して改変する改変手段とを備える。

- [0011] 前記改変手段は、キャッシュエントリーに対してタグとしてのアドレスを設定し、有効なデータを保持していない状態でバリッドフラグをセットする構成としてもよい。
- [0012] この構成によれば、配列などのデータを書き込むためのキャッシュエントリーをキャッシュメモリ上に確保することができ、しかも、上書きされる運命にあるデータをメモリからキャッシュメモリへ無駄にロードすることを防止することができる。
- [0013] また、前記改変手段は、プロセッサからの指定されたキャッシュエントリーに対してライトバックされていない書き換えられたデータを保持している状態でダーティフラグをリセットする。
- [0014] この構成によれば、テンポラリーなワークデータ等破棄される運命にあるデータを保持するキャッシュエントリーから無駄にライトバックすることを防止することができる。
- [0015] ここで、前記キャッシュメモリは、さらに、プロセッサから指定されたアドレス範囲を保持する保持手段と、保持されたアドレス範囲に属するデータを保持するキャッシュエントリーを特定する特定手段とを備え、前記改変手段は、特定されたキャッシュエントリーに対して前記バリッドフラグおよびダーティフラグの少なくとも一方を改変する構成としてもよい。
- [0016] この構成によれば、プロセッサから指定されたアドレス範囲で、キャッシュメモリへの無駄なデータのロード又はメモリへの無駄なライトバックを防止することができる。
- [0017] ここで、前記特定手段は、前記アドレス範囲の先頭アドレスがラインデータの途中を指す場合、当該先頭アドレスを、前記アドレス範囲に含まれる先頭のラインを指すスタートラインアドレスに変換する第1変換手段と、前記アドレス範囲の末尾アドレスがラインデータの途中を指す場合、当該末尾アドレスを、前記アドレス範囲に含まれる末尾のラインを指すエンドラインアドレスに変換する第2変換手段と、前記スタートラインアドレスからエンドラインアドレスまでの各ラインアドレスに対応するデータを保持するキャッシュエントリーがあるか否かを判定する判定手段と備える構成としてもよい。
- [0018] この構成によれば、プロセッサは、前記アドレス範囲として、キャッシュメモリのラインサイズ及びライン境界のアドレスとは無関係に任意のアドレスから任意のアドレス(又は任意のサイズ)を指定することができる。つまり、プロセッサにおいてキャッシュメモリのラインサイズ及びライン境界のアドレスを管理する必要がないので、キャッシュメモリ

管理のための負荷を解消することができる。

[0019] ここで、前記改変手段は、ダーティフラグのリセット指示付きメモリアクセス命令が実行されたことを検出する命令検出手段と、当該命令によってアクセスがなされたキャッシュエントリーに対してダーティフラグをリセットするフラグ書き換え手段とを備える構成としてもよい。

[0020] ここで、前記改変手段は、バリッドフラグのリセット指示付きメモリアクセス命令が実行されたことを検出する命令検出手段と、当該命令によってアクセスがなされたキャッシュエントリーに対してバリッドフラグをリセットするフラグ書き換え手段とを備える構成としてもよい。

また、本発明のキャッシュメモリの制御方法についても上記と同様の手段、作用を有する。

[0021] 本発明のキャッシュメモリによれば、配列などのデータを書き込むためのキャッシュエントリーをキャッシュメモリ上に確保することができ、しかも、上書きされる運命にあるデータをメモリからキャッシュメモリへ無駄にロードすることを防止することができる。

また、テンポラリーなワークデータ等破棄される運命にあるデータを保持するキャッシュエントリーから無駄にライトバックすることを防止することができる。

[0022] さらに、プロセッサは、前記アドレス範囲として、キャッシュメモリのラインサイズ及びライン境界のアドレスとは無関係に任意のアドレスから任意のアドレス(又は任意のサイズ)を指定することができる。つまり、プロセッサにおいてキャッシュメモリのラインサイズ及びライン境界のアドレスを管理する必要がないので、キャッシュメモリ管理のための負荷を解消することができる。

発明の効果

[0023] 本発明のキャッシュメモリによれば、プロセッサの指示に基づいて、配列などのデータを書き込むためのキャッシュエントリーをキャッシュメモリ上に確保することができ、しかも、上書きされる運命にあるデータをメモリからキャッシュメモリへ無駄にロードすることを防止することができる。また、テンポラリーなワークデータ等破棄される運命にあるデータを保持するキャッシュエントリーから無駄にライトバックすることを防止することができる。

図面の簡単な説明

- [0024] [図1]本発明の実施の形態1におけるプロセッサ、キャッシュメモリ、メモリを含むシステムの概略構成を示すブロック図である。
- [図2]キャッシュメモリの構成例を示すブロック図である。
- [図3]キャッシュエントリーの詳細なビット構成を示す図である。
- [図4]制御部の構成を示すブロック図である。
- [図5]フラグ改変部の構成例を示すブロック図である。
- [図6] (a) スタートアドレスレジスタにスタートアドレスを書き込む命令の一例を示す。(b) サイズレジスタにサイズを書き込む命令の一例を示す。(c) コマンドレジスタにコマンドを書き込む命令の一例を示す。(d) コマンドの一例を示す。
- [図7] スタートアライナ及びエンドアライナの説明図である。
- [図8] フラグ書換部407におけるフラグ改変処理の一例を示すフローチャートである。
- [図9] 本発明の実施の形態2におけるキャッシュメモリの構成を示すブロック図である。
- [図10] キャッシュエントリーのビット構成を示す。
- [図11] 制御部の構成を示すブロック図である。
- [図12] フラグ更新部による使用フラグUの更新例を示す。
- [図13] (a) ウィークフラグが存在しない場合にキャッシュエントリーがリプレースされる様子を示す図である。(b) リプレース処理におけるウィークフラグWの役割を示す説明図である。
- [図14] リプレース部におけるUフラグ更新処理を示すフローチャートである。
- [図15] リプレース部におけるリプレース処理を示すフローチャートである。
- [図16] フラグ改変部の構成を示すブロック図である。
- [図17] コマンドフォーマットの一例を示す。
- [図18] フラグ書換部におけるWフラグ設定処理の一例を示すフローチャートである。

符号の説明

- [0025]
- 1 プロセッサ
 - 2 メモリ
 - 3 キャッシュメモリ

- 20 アドレスレジスタ
- 21 メモリI/F
- 30 デコーダ
- 31a〜31d ウェイ
- 32a〜32d 比較器
- 33a〜33d アンド回路
- 34 オア回路
- 35 セレクタ
- 36 セレクタ
- 37 デマルチプレクサ
- 38 制御部
- 39 フラグ更新部
- 40 リプレース部
- 41 フラグ改変部
- 131a〜131d ウェイ
- 138 制御部
- 139 フラグ更新部
- 140 リプレース部
- 141 フラグ改変部
- 401 コマンドレジスタ
- 402 スタートアドレスレジスタ
- 403 サイズレジスタ
- 404 加算器
- 405a スタートアライナ
- 405b スタートアライナ
- 406a エンドアライナ
- 406b エンドアライナ
- 407 フラグ書換部

407a フラグ書換部

発明を実施するための最良の形態

[0026] (実施の形態1)

<全体構成>

図1は、本発明の実施の形態1におけるプロセッサ1、キャッシュメモリ3、メモリ2を含むシステムの概略構成を示すブロック図である。同図のように、本発明のキャッシュメモリ3は、プロセッサ1およびメモリ2を有するシステムに備えられる。

キャッシュメモリ3は、キャッシュエントリー毎にキャッシュエントリーが有効であるか否かを示すバリッドフラグVと、当該キャッシュエントリーに対して書き込みがなされたか否かを示すダーティフラグDとを有し、プロセッサ1により指定されたアドレスのデータを保持するキャッシュエントリーに対してデータをライトバックすることなくダーティフラグDを1から0(ダーティでない)に改変し、また、プロセッサ1により指定されたアドレスに対応するキャッシュエントリーを確保してメモリからデータをロードすることなくバリッドフラグVを1(有効)に改変するよう構成されている。

ダーティフラグDを1から0に改変するのは、最終的には破棄されるテンポラリーなワークデータを保持するキャッシュエントリーから無駄にライトバックすることを防止するためである。また、キャッシュエントリーを確保してメモリからデータをロードすることなくバリッドフラグVを1(有効)に改変することにより、配列など書き込みをするためのキャッシュエントリーを予め確保するためである。

[0027] <キャッシュメモリの構成>

以下、キャッシュメモリ3の具体例として、4ウェイ・セット・アソシエイティブ方式のキャッシュメモリに本発明を適用した場合の構成について説明する。

図2は、キャッシュメモリ3の構成例を示すブロック図である。同図のように、キャッシュメモリ3は、アドレスレジスタ20、メモリI/F21、デコーダ30、4つのウェイ31a〜31d(以下ウェイ0〜3と略す)、4つの比較器32a〜32d、4つのアンド回路33a〜33d、オア回路34、セレクト35、36、デマルチプレクサ37、制御部38を備える。

[0028] アドレスレジスタ20は、メモリ2へのアクセスアドレスを保持するレジスタである。このアクセスアドレスは32ビットであるものとする。同図に示すように、アクセスアドレスは、

最上位ビットから順に、21ビットのタグアドレス、4ビットのセットインデックス(図中のSI)、5ビットのワードインデックス(図中のWI)を含む。ここで、タグアドレスはウェイにマッピングされるメモリ中の領域(そのサイズはセット数×ブロックである)を指す。この領域のサイズは、タグアドレスよりも下位のアドレスビット(A10〜A0)で定まるサイズつまり2kバイトであり、1つのウェイのサイズでもある。セットインデックス(SI)はウェイ0〜3に跨る複数セットの1つを指す。このセット数は、セットインデックスが4ビットなので16セットある。タグアドレスおよびセットインデックスで特定されるキャッシュエントリーは、リプレース単位であり、キャッシュメモリに格納されている場合はラインデータ又はラインと呼ばれる。ラインデータのサイズは、セットインデックスよりも下位のアドレスビットで定まるサイズつまり128バイトである。1ワードを4バイトとすると、1ラインデータは32ワードである。ワードインデックス(WI)は、ラインデータを構成する複数ワード中の1ワードを指す。アドレスレジスタ20中の最下位2ビット(A1、A0)は、ワードアクセス時には無視される。

[0029] メモリI/F21は、キャッシュメモリ3からメモリ2へのデータのライトバックや、メモリ2からキャッシュメモリ3へのデータのロード等、キャッシュメモリ3からメモリ2をアクセスするためのI/Fである。

[0030] デコーダ30は、セットインデックスの4ビットをデコードし、4つのウェイ0〜3に跨る16セット中の1つを選択する。

[0031] 4つのウェイ0〜3は、同じ構成を有数する4つのウェイであり、4×2kバイトの容量を有する。各ウェイは、16個のキャッシュエントリーを有する。

[0032] 図3に1つのキャッシュエントリーにおける詳細なビット構成を示す。同図のように、1つのキャッシュエントリーは、バリッドフラグV0〜V3、21ビットのタグ、128バイトのラインデータ、ダーティフラグD0〜D3を有する。

[0033] タグは21ビットのタグアドレスのコピーである。

ラインデータは、タグアドレスおよびセットインデックスにより特定されるブロック中の128バイトデータのコピーであり、32バイトの4つのサブラインからなる。

[0034] バリッドフラグV0〜V3は、4つのサブラインに対応し、サブラインが有効か否かを示す。

- [0035] ダーティフラグD0〜D3は、4つのサブラインに対応し、そのサブラインにプロセッサから書き込みがあったか否か、つまりサブライン中にキャッシュされたデータが存在するが書き込みによりメモリ中のデータと異なるためメモリに書き戻すことが必要か否かを示す。
- [0036] 比較器32aは、アドレスレジスタ20中のタグアドレスと、セットインデックスにより選択されたセットに含まれる4つのタグ中のウェイ0のタグとが一致するか否かを比較する。比較器32b〜32cについても、ウェイ31b〜31dに対応すること以外は同様である。
- [0037] アンド回路33aは、バリッドフラグと比較器32aの比較結果とが一致するか否かを比較する。この比較結果をh0とする。比較結果h0が1である場合は、アドレスレジスタ20中のタグアドレスおよびセットインデックスに対応するラインデータが存在すること、つまりウェイ0においてヒットしたことを意味する。比較結果h0が0である場合は、ミスヒットしたことを意味する。アンド回路33b〜33dについても、ウェイ31b〜31dに対応すること以外は同様である。その比較結果h1〜h3は、ウェイ1〜3でヒットしたかミスしたかを意味する。
- [0038] オア回路34は、比較結果h0〜h3のオアをとる。このオアの結果をhitとする。hitは、キャッシュメモリにヒットしたか否かを示す。
- [0039] セレクタ35は、選択されたセットにおけるウェイ0〜3のラインデータのうち、ヒットしたウェイのラインデータを選択する。
- [0040] セレクタ36は、セレクタ35により選択された32ワードのラインデータのうち、ワードインデックスに示される1ワードを選択する。
- [0041] デマルチプレクサ37は、キャッシュエントリにデータを書き込む際に、ウェイ0〜3の1つに書き込みデータを出力する。この書き込みデータはワード単位でよい。
- [0042] 制御部38は、キャッシュメモリ3の全体の制御を行う。特に、プロセッサからのコマンド及びアドレス指定に従って、Vフラグの改変と、Dフラグの改変とを行う。
- [0043] <制御部の構成>

図4は、制御部38の構成を示すブロック図である。同図のように、制御部38は、フラグ更新部39、リプレース部40、フラグ改変部41とを含む。

フラグ更新部39は、従来技術と同様にキャッシュメモリの状態を反映するようVフラ

グ、Dフラグの更新を行う。

リプレース部40は、従来技術と同様にキャッシュエントリーのリプレースを行う。

フラグ改変部41は、プロセッサ1からのコマンド及びアドレス指定に応じて、Vフラグの改変とDフラグの改変とを行う。このコマンドには、Vフラグ設定コマンドとDフラグ設定コマンドとがある。

[0044] <フラグ改変部の構成>

図5は、フラグ改変部41の構成例を示すブロック図である。同図のようにフラグ改変部41は、コマンドレジスタ401、スタートアドレスレジスタ402、サイズレジスタ403、加算器404、スタートアライナ405a、405b、エンドアライナ406a、406b、フラグ書換部407を備える。

[0045] コマンドレジスタ401は、プロセッサ1から直接アクセス可能なレジスタであり、プロセッサ1により書き込まれたコマンドを保持する。図6(c)に、コマンドレジスタ401にコマンドを書き込む命令の一例を示す。この命令は、通常の転送命令(mov命令)であり、ソースオペランドとしてコマンドを、デスティネーションオペランドとしてコマンドレジスタ(CR)401を指定している。図6(d)に、コマンドフォーマットの一例を示す。このコマンドフォーマットは、Vフラグ設定コマンド用の2ビットのフィールドと、Dフラグ設定コマンド用の2ビットのフィールドと、サブライン指定用の4ビットのフィールドとを含む。

[0046] 例えば、Vフラグ設定コマンドは、「10」であればV=0に設定することを指示し(Vリセットコマンド)、「11」であればV=1に設定することを指示し(Vセットコマンド)、「00」であれば無効(ノーオペレーション)を意味する。Dフラグ設定コマンドも同様である。サブライン指定フィールドは、サブラインを個別に指定するためのフィールドであり、個別指定がない場合はライン(全サブライン)指定であるものとする。

[0047] スタートアドレスレジスタ402は、プロセッサ1から直接アクセス可能なレジスタであり、プロセッサ1により書き込まれたスタートアドレスを保持する。このスタートアドレスはCフラグを設定すべきアドレス範囲の開始位置を示す。図6(a)に、スタートアドレスレジスタ402にスタートアドレスを書き込む命令の一例を示す。この命令も、図6(c)と同様に通常の転送命令(mov命令)である。

[0048] サイズレジスタ403は、プロセッサ1から直接アクセス可能なレジスタであり、プロセッ

サ1により書き込まれたサイズを保持する。このサイズは、スタートアドレスからのアドレス範囲を示す。図6(b)に、サイズレジスタ403にサイズを書き込む命令の一例を示す。この命令も、図6(c)と同様に通常の転送命令(mov命令)である。なお、サイズの単位は、バイト数であっても、ライン数(キャッシュエントリー数)であってもよく、予め定められた単位であればよい。

- [0049] 加算器404は、スタートアドレスレジスタ402に保持されたスタートアドレスとサイズレジスタ403に保持されたサイズとを加算する。加算結果は、アドレス範囲の終了位置を指すエンドアドレスである。加算器404は、サイズがバイト数指定の場合はバイトアドレスとして加算し、サイズがライン数指定の場合はラインアドレスとして加算すればよい。
- [0050] スタートアライナ405a、405bは、スタートアドレスをライン境界の位置に調整する。スタートアライナ405aはエンドアドレスの方向に、405bはエンドアドレスとは反対の方向に調整する。この調整によりプロセッサ1はラインサイズ及びライン境界とは無関係に任意のアドレスをスタートアドレスとして指定することができる。
- [0051] エンドアライナ406a、406bは、エンドアドレスをライン境界の位置に調整する。エンドアライナ4065aはスタートアドレスの方向に、406bはスタートアドレスとは反対の方向に調整する。この調整によりプロセッサ1はラインサイズ及びライン境界とは無関係に任意の大きさを上記サイズとして指定することができる。
- [0052] 図7に、スタートアライナ405a、405b及びエンドアライナ406a、406bの説明図を示す。同図において、プロセッサ1から指定されたスタートアドレスはラインNの途中の任意の位置を指す。スタートアライナ405aは、次のライン(N+1)の先頭を指すよう調整し、調整後のアドレスをアラインスタートアドレスaとして出力する。スタートアライナ405bは、スタートアドレスのデータを含むラインNの先頭を指すよう調整し、調整後のアドレスをアラインスタートアドレスbとして出力する。アラインスタートアドレスが指すラインをスタートラインと呼ぶ。
- [0053] また、エンドアドレスはラインMの途中の任意の位置を指す。エンドアライナ406aは、直前のライン(M-1)の先頭を指すよう調整し、調整後のアドレスをアラインエンドアドレスaとして出力する。エンドアライナ406bは、エンドアドレスのデータを含むライン

Mの先頭を指すよう調整し、調整後のアドレスをアラインエンドアドレスbとして出力する。アラインエンドアドレスが指すラインをエンドラインと呼ぶ。

[0054] 同図のように、スタートアライナ405a及びエンドアライナ406aはライン単位で内側アラインを行う。スタートアライナ405b及びエンドアライナ406bはライン単位で外側アラインを行う。さらに、ライン単位の外側アラインの後、さらに、サブライン単位の外側アラインと内側アラインが可能である。

[0055] フラグ書換部407は、スタートラインからエンドラインまで、コマンドに従ってVフラグ又はDフラグの値を設定する。その際、スタートライン及びエンドラインが内側アラインか外側アラインかはコマンドに応じて選択される。

[0056] <フラグ改変処理>

図8は、フラグ書換部407におけるフラグ改変処理の一例を示すフローチャートである。同図では、DリセットコマンドとVセットコマンドとを示している。

同図において、コマンドレジスタ401にDリセットコマンドが書き込まれている場合(S80)、フラグ書換部407は、スタートアライナ405a、405b、エンドアライナ406a、406bの出力の中から内側アラインによるスタートラインとエンドラインとを選択する(S81)。ここで内側アラインを選択しているのは、ラインNとラインMにおけるスタートアドレスとエンドアドレスの外側のデータは、プロセッサ1により破棄されるとは限らないからである。

[0057] さらに、フラグ書換部407は、スタートラインからエンドラインまでの各ラインアドレスを順に出力しながらループ1の処理(S82〜S86)を行う。フラグ書換部407は、各ラインについて同じ処理を行うので、ここでは1ライン分の処理について説明する。

[0058] すなわち、フラグ書換部407は、キャッシュメモリ3がプロセッサ1からアクセスされていない間に、ラインアドレスをアドレスレジスタ20に出力し(S83)、アドレスレジスタ20のタグアドレスとキャッシュエントリーのタグとを比較器32a〜32dに比較させ、ヒットするかどうかを判定する(S84)。さらにフラグ書換部407は、ヒットした場合には、ヒットしたキャッシュエントリーに対してD0〜D3フラグを0にリセットし(S85)、ミスヒットした場合には、キャッシュメモリにエントリーされていないのでなにもしない。

[0059] このように、スタートラインからエンドラインまでの各ラインについて、キャッシュメモリ

3にエントリーされている場合には、D0〜D3フラグに0が設定される。フラグ書換部407は、ループ1の終了後、コマンドレジスタ401のコマンドをクリアする(S98)。これにより、テンポラリーなワークデータ等破棄される運命にあるデータを保持するキャッシュエントリーから無駄にライトバックすることを防止することができる。

[0060] また、図8において、コマンドレジスタ401にVセットコマンドが書き込まれている場合(S87)、フラグ書換部407は、スタートアライナ405a、405b、エンドアライナ406a、406bの出力の中から外側アラインによるスタートラインとエンドラインとを選択する(S88)。ここで外側アラインを選択しているのは、内側アラインではスタートアドレスからエンドアドレスまでのサイズを確保できないからである。

[0061] さらに、フラグ書換部407は、スタートラインからエンドラインまでの各ラインアドレスを順に出力しながらループ2の処理(S89〜S97)を行う。フラグ書換部407は、各ラインについて同じ処理を行うので、ここでは1ライン分の処理について説明する。

[0062] すなわち、フラグ書換部407は、キャッシュメモリ3がプロセッサ1からアクセスされていない間に、ラインアドレスをアドレスレジスタ20に出力し(S90)、アドレスレジスタ20のタグアドレスとキャッシュエントリーのタグとを比較器32a〜32dに比較させ、ヒットするかどうかを判定する(S91)。さらにフラグ書換部407は、ヒットしなかった場合には、当該ラインアドレスに対応するセット内の4つのウェイからLRU方式でリプレース対象のウェイを1つ選択し(S92)、選択されたウェイのダーティフラグD0〜D3の論理和が1であるか否かを判定する(S93)。この論理和が1、つまりダーティであると判定された場合、フラグ書換部407は、ダーティなサブラインのみをライトバックする(S94)。上記の論理和が1でないつまりダーティでないと判定された場合、又はダーティなサブラインのライトバックの後に、フラグ書換部407は、キャッシュエントリーにメモリからデータをロードすることなく、キャッシュエントリーにラインアドレスをタグとして設定し(S95)、バリッドフラグV0〜V3を1にセットする(S96)。このようにして、有効なデータを保持していないがV0〜V3=1と設定された1ライン分のキャッシュエントリーが確保される。

[0063] さらに、フラグ書換部407は、ループ1の終了後、コマンドレジスタ401のコマンドをクリアする(S98)。

- [0064] これにより、配列などのデータを書き込み用のキャッシュエントリーをキャッシュメモリ上に確保し、かつ、無駄なライトバックを防止することができる。
- [0065] 以上説明してきたように、本実施の形態におけるキャッシュメモリによれば、Dリセットコマンドにより、破棄される運命にあるデータを保持するキャッシュエントリーから無駄にライトバックすることを防止することができる。また、Vセットコマンドにより、配列などのデータを書き込み用のキャッシュエントリーをキャッシュメモリ上に確保し、かつ、無駄なライトバックを防止することができる。
- [0066] なお、上記実施の形態では、DリセットコマンドとVセットコマンドについて説明したが、Vリセットコマンドについては、図8に示したS87～S98において、ステップS95を削除し、ステップS96においてV0～V3を0にリセットすることにより実現することができる。これによれば、無駄なライトバックを防止してキャッシュエントリーを開放することができる。
- [0067] また、Dセットコマンドは、図8のS80～S86と同様に実行可能ではあるが、プロセッサ1からストア命令によりデータを書き込めばDフラグがセットされること、Dフラグのセットによりライトバック動作が発生することを考えれば、あまり実用的とはいえない。しかし、キャッシュメモリのテスト動作や性能の評価や検証等に利用することができる。
- [0068] <変形例>

なお、本発明のキャッシュメモリは、上記の実施の形態の構成に限るものではなく、種々の変形が可能である。以下、変形例のいくつかについて説明する。

(1) 上記実施の形態では、V0～V3フラグ、D0～D3フラグを同時にセット又はリセットしているが、サブ単位にセット又はリセットするようにしてもよい。

例えば、スタートラインとエンドラインのみをサブライン単位で処理する場合、フラグ書換部407は、外側アライン(ライン)されたスタートアドレスとエンドアドレスを選択し、さらに外側アライン(サブライン)または内側アライン(サブライン)によりスタートラインのサブラインアドレスと、エンドラインのサブラインアドレスとを算出し、スタートラインとエンドラインのみをサブライン単位で処理すればよい。また、プロセッサ1は、コマンド中のサブライン指定フィールドにおいてその旨を指定すればよい。

例えば、プロセッサ1がサイズレジスタ403のサイズを0に指定し、コマンド中のサブ

ライン指定フィールドにおいて特定のサブラインを指定した場合に、フラグ書換部407は、指定されたサブラインのみを対象に処理をすればよい。

(2) また、フラグ改変部41は、ダーティフラグのリセット指示付きメモリアクセス命令が実行されたこと検出する命令検出部と、当該命令によってアクセスがなされたキャッシュエントリーに対してダーティフラグをリセットするフラグ書き換え手部とを備える構成としてもよい。

さらに、命令検出部は、バリッドフラグのリセット指示付きメモリアクセス命令が実行されたこと検出し、フラグ書き換え手部は、当該命令によってアクセスがなされたキャッシュエントリーに対してバリッドフラグをリセットする構成としてもよい。

(3) 上記実施の形態では、4ウェイ・セット・アソシエイティブのキャッシュメモリを例に説明したが、ウェイ数は、いくつでもよい。また、上記実施の形態では、セット数が16である例を説明したが、セット数はいくつでもよい。

(4) 上記実施の形態では、セット・アソシエイティブのキャッシュメモリを例に説明したが、フル・アソシエイティブ方式のキャッシュメモリであってもよい。

(5) 上記実施の形態では、サブラインのサイズをラインのサイズの1/4としているが、1/2、1/8、1/16等他のサイズでもよい。その場合、各キャッシュエントリーは、サブラインと同数のバリッドフラグおよびダーティフラグをそれぞれ保持すればよい。

[0069] (実施の形態2)

実施の形態1では、Vフラグの改変とDフラグの改変とを行う構成について説明したが、本実施の形態では、これ以上使用(書き込み及び読み出し)するかしないかを示すW(ウィーク)フラグを有するキャッシュメモリにおいてWフラグを改変する構成について説明する。

[0070] <キャッシュメモリの構成>

図9は、本発明の実施の形態2におけるキャッシュメモリの構成を示すブロック図である。同図のキャッシュメモリは、図2の構成と比較して、ウェイ31a〜31dの代わりにウェイ131a〜131dを備える点と、制御部38の代わりに制御部138を備える点とが異なっている。以下、同じ点は説明を省略して、異なる点を中心に説明する。

[0071] ウェイ131aは、ウェイ31aと比べて、各キャッシュエントリー中にWフラグ及びUフラ

グが追加されている点が異なる。ウェイ131b〜131dも同様である。

[0072] 図10に、キャッシュエントリーのビット構成を示す。1つのキャッシュエントリーは、バリッドフラグV0〜V3、21ビットのタグ、128バイトのラインデータ、ウィークフラグW、使用フラグU及びダーティフラグD0〜D3を保持する。

[0073] このうち、ウィークフラグWは、プロセッサからのアクセスに関しては、これ以上使用するか否かを意味し、キャッシュメモリにおけるリプレース制御に関しては、他のキャッシュエントリーよりも真っ先に追い出してもよい最弱のリプレース対象を意味する。このように、ウィークフラグWは二つの意味を有することから、クリーニング処理とリプレース処理との2つの処理で参照される。

[0074] 使用フラグUは、そのキャッシュエントリーにアクセスがあったか否かを示し、LRU方式におけるミスヒットによるリプレースに際して4つのウェイのキャッシュエントリー間におけるアクセス順序データの代わりに用いられる。より正確には、使用フラグUの1は、アクセスがあったことを、0はないことを意味する。ただし、1つのセット内の4つウェイの使用フラグが全て1になれば、0にリセットされる。別言すれば、使用フラグUは、アクセスされた時期が古いか新しいか2つの相対的な状態を示す。つまり、使用フラグUが1のキャッシュエントリーは、使用フラグが0のキャッシュエントリーよりも新しくアクセスされたことを意味する。

[0075] 制御部138は、制御部38と比べて、Wフラグを設定する点と、LRU方式におけるアクセス順序情報の代わりに使用フラグUを用いる点とが異なる。

[0076] <制御部の構成>

図11は、制御部138の構成を示すブロック図である。同図の制御部138は、制御部38と比較して、フラグ更新部39、リプレース部40、フラグ改変部41の代わりにフラグ更新部139、リプレース部140、フラグ改変部141を備える点異なる。

[0077] フラグ更新部139は、フラグ更新部39と同様にVフラグ、Dフラグを更新することに加えて、キャッシュメモリがアクセスされたときに使用フラグUの更新処理を行う。

[0078] リプレース部140は、通常のLRU方式ではなく、使用フラグUをアクセス順序とする擬似的なLRU方式によりリプレースを行う。ただし、リプレース処理に際してW=1のキャッシュエントリーは真っ先にリプレース対象として選択する。

[0079] フラグ改変部141は、プロセッサ1からのコマンドに応じてウィークフラグWを設定する。プロセッサ1は、もはや使用(書き込み及び読み出し)をしないキャッシュエントリーについてウィークフラグの設定を指示するコマンドをキャッシュメモリ3に対して発行する。W=1のキャッシュエントリーは、キャッシュミス時には使用フラグUの値に関わらず、真っ先にリプレース対象となる。また、W=1のキャッシュエントリーがダーティであればクリーニング処理の対象となる。

[0080] <使用フラグUの説明>

図12は、フラグ更新部39による使用フラグUの更新例を示す。同図の上段、中断、下段は、ウェイ0〜3に跨るセットNを構成する4つのキャッシュエントリーを示している。4つのキャッシュエントリー右端の1又は0は、それぞれ使用フラグの値である。この4つの使用フラグUをU0〜U3と記す。

[0081] 同図上段では(U0〜U3)=(1, 0, 1, 0)であるので、ウェイ0、2のキャッシュエントリーはアクセスがあったことを、ウェイ1、3のキャッシュエントリーはアクセスがないことを意味する。

[0082] この状態で、メモリアクセスがセットN内のウェイ1のキャッシュエントリーにヒットした場合、同図中段に示すように、(U0〜U3)=(1, 1, 1, 0)に更新される。つまり、実線に示すようにウェイ1の使用フラグU1が0から1に更新される。

[0083] さらに、同図中段の状態で、メモリアクセスがセットN内のウェイ3のキャッシュエントリーにヒットした場合、同図下断に示すように、(U0〜U3)=(0, 0, 0, 1)に更新される。つまり、実線に示すようにウェイ3の使用フラグU1が0から1に更新される。加えて、破線に示すようにウェイ3以外の使用フラグU0〜U2が1から0に更新される。これにより、ウェイ3のキャッシュエントリーが、ウェイ0〜2の各キャッシュエントリーよりも新しくアクセスされたことを意味することになる。

[0084] リプレース部140は、キャッシュミス時にW=1のキャッシュエントリーが存在しなければ、使用フラグに基づいてリプレース対象のキャッシュエントリーを決定してリプレースを行う。例えば、フラグ更新部39は、図5上段では、ウェイ1とウェイ3の何れかをリプレース対象と決定し、図5中段ではウェイ3をリプレース対象と決定し、図5下段ではウェイ0〜2の何れかをリプレース対象と決定する。

[0085] <ウィークフラグWの説明>

図13(a)ウィークフラグが存在しないと仮定した場合の比較例であり、キャッシュエントリーがリプレイスされる様子を示す図である。同図においても、図12と同様にウェイ0〜3に跨るセットNを構成する4つのキャッシュエントリーを示している。、4つのキャッシュエントリー右端の1又は0は、それぞれ使用フラグの値である。また、データEのみアクセス頻度の低いデータを、データA、B、C、Dはアクセス頻度の高いデータとする。

[0086] 同図(a)の第1段目の状態で、プロセッサ1がデータEにアクセスすると、キャッシュミスが発生する。このキャッシュミスにより、例えば、U=0のキャッシュエントリーの中からアクセス頻度の高いデータCのキャッシュエントリーがアクセス頻度の低いデータEにリプレイスされ、第2段目の状態となる。

[0087] 第2段目の状態で、プロセッサ1がデータCにアクセスすると、キャッシュミスが発生する。このキャッシュミスにより、U=0のキャッシュエントリーであるアクセス頻度の高いデータDのキャッシュエントリーがアクセス頻度の高いデータCにリプレイスされ、第3段目の状態となる。

[0088] 第3段目の状態で、プロセッサ1がデータDにアクセスすると、キャッシュミスが発生する。このキャッシュミスにより、例えば、アクセス頻度の高いデータCのキャッシュエントリーがアクセス頻度の高いデータDにリプレイスされ、第3段目の状態となる。

[0089] 同様に、第4段目でも、使用頻度の低いデータEはリプレイス対象として選択されないで、キャッシュメモリーに残っている。

[0090] 第5段目の状態で、使用頻度の低いデータEは最も古い(U=0)ことから、リプレイス対象として選択されて、追い出される。

[0091] このように、擬似LRU方式において(通常のLRU方式においても)、アクセス頻度の低いデータEによって、4ウェイの場合は最悪4回のキャッシュミスを誘発する場合がある。

図13(b)は、リプレイス処理におけるウィークフラグWの役割を示す説明図である。

[0092] 同図(b)の第1段目の状態(同図(a)の第1段目と同じ)で、プロセッサ1がデータEにアクセスすると、キャッシュミスが発生する。このキャッシュミスにより、例えば、U=0

のキャッシュエントリーの中からアクセス頻度の高いデータCのキャッシュエントリーがアクセス頻度の低いデータEにリプレイスされる。このとき、プロセッサ1は、データEのキャッシュエントリーにウィークフラグWを1に設定するものとする。これにより、次のキャッシュミス時にデータEのキャッシュエントリーが真っ先に追い出され、第2段目の状態となる。

[0093] 第2段目の状態で、プロセッサ1がデータCにアクセスすると、キャッシュミスが発生する。このキャッシュミスにより、 $W=1$ のキャッシュエントリーであるアクセス頻度の低いデータEのキャッシュエントリーがリプレイス対象として選択され、アクセス頻度の高いデータCにリプレイスされ、第3段目の状態となる。

[0094] このように、ウィークフラグWを設けることにより、アクセス頻度の低いデータによるキャッシュミスの誘発を低減することができる。

[0095] <Uフラグ更新処理>

図14は、リプレイス部140におけるUフラグ更新処理を示すフローチャートである。同図では、バリッドフラグが0(無効)であるキャッシュエントリーの使用フラグUは0に初期化されているものとする。

同図において、リプレイス部140は、キャッシュヒットしたとき(ステップS61)、セットインデックスにより選択されたセットにおけるヒットしたウェイの使用フラグUを1にセットし(ステップS62)、そのセット内の他のウェイの使用フラグUを読み出し(ステップS63)、読み出した使用フラグUが全て1であるか否かを判定し(ステップS64)、全て1でなければ終了し、全て1であれば他のウェイの全ての使用フラグUを0にリセットする(ステップS65)。

このようにしてリプレイス部140は、図12、図13(a)(b)に示した更新例のように、使用フラグUを更新する。

[0096] <リプレイス処理>

図15は、リプレイス部140におけるリプレイス処理を示すフローチャートである。同図においてリプレイス部140は、メモリアクセスがミスしたとき(ステップS91)、セットインデックスにより選択されたセットにおける、4つウェイの使用フラグU及びウィークフラグWを読み出し(ステップS92)、 $W=1$ のウェイが存在するか否かを判定する(ステッ

プS93)。W=1のウェイが存在しないと判定された場合、U=0のウェイを1つ選択する(ステップS94)。このとき、使用フラグUが0になっているウェイが複数存在する場合は、リプレース部140はランダムに1つを選択する。また、W=1のウェイが存在すると判定された場合、Uフラグの値に関わらずW=1のウェイを1つ選択する(ステップS95)。このとき、ウィークフラグWが1になっているウェイが複数存在する場合は、リプレース部140はランダムに1つを選択する。

[0097] さらに、リプレース部140は、当該セットにおける選択されたウェイのキャッシュエントリーを対象にリプレースし(ステップS96)、リプレース後に当該キャッシュエントリーの使用フラグUを1に、ウィークフラグWを0初期化する(ステップS97)。なお、このときバリッドフラグV、ダーティフラグDは、それぞれ1、0に初期化される。

このように、W=1のウェイが存在しない場合、リプレース対象は、使用フラグUが0のキャッシュエントリーの中から1つ選択される。

また、W=1のウェイが存在する場合、リプレース対象は、使用フラグUが0であると1であるとを問わず、W=1のウェイのキャッシュエントリーから1つ選択される。これにより図13(a)(b)に示したように、アクセス頻度の低いデータがキャッシュメモリに残ることによるキャッシュミスの誘発を低減することができる。

[0098] <フラグ改変部141の構成>

図16は、フラグ改変部141の構成を示すブロック図である。同図の構成は、図5に示したフラグ改変部41と比較して、フラグ書換部407の代わりにフラグ書換部407aを備える点が異なっている。

[0099] フラグ書換部407aは、フラグ書換部407の機能に加えて、Wフラグの改変と使用フラグUの改変とを行う点が異なっている。そのため、コマンドレジスタ401には、Wフラグの設定を指示するWコマンド、Uフラグの設定を指示するUコマンドがプロセッサ1によって設定可能になっている。図17に、これらのコマンドフォーマットの一例を示す。同図のコマンドフォーマットは、図6(d)に示したコマンドフォーマットに対してWコマンド及びUコマンドのフィールドが追加されている。Wコマンド及びUコマンドの内容はDコマンドやVコマンドと同様である。

[0100] <Wフラグ設定処理>

図18は、フラグ書換部407aにおけるWフラグ設定処理の一例を示すフローチャートである。

フラグ書換部407aは、コマンドレジスタ401にWフラグ設定コマンドが保持されている場合、スタートラインからエンドラインまでの各ラインアドレスを順に出力しながらループ1の処理(S82〜S86)を行う。フラグ書換部407aは、各ラインについて同じ処理を行うので、ここでは1ライン分の処理について説明する。

すなわち、フラグ書換部407aは、キャッシュメモリ3がプロセッサ1からアクセスされていない間に、ラインアドレスをアドレスレジスタ20に出力し(S83)、アドレスレジスタ20のタグアドレスとキャッシュエントリーのタグとを比較器32a〜32dに比較させ、ヒットするかどうかを判定する(S84)。さらにフラグ書換部407aは、ヒットした場合には、ヒットしたキャッシュエントリーに対してWフラグを1にセットし(S85)、ミスヒットした場合には、キャッシュメモリにエントリーされていないのでなにもしない。

これにより、スタートラインからエンドラインまでの各ラインについて、キャッシュメモリ3にエントリーされている場合には、Wフラグが1に設定される。

[0101] <Uフラグ設定処理>

フラグ書換部407aは、コマンドレジスタ401にUフラグ設定コマンドが保持されている場合、当該コマンドに従ってUフラグを設定する。この処理は、図18においてWフラグをUフラグと読み替えることにより、Wフラグ設定処理と全く同様に実行される。

[0102] 以上説明してきたように、本実施の形態におけるキャッシュメモリによれば、Wフラグを設定することにより、W=1のキャッシュエントリーをキャッシュミス時には最古のキャッシュエントリとして真っ先にリプレース対象とすることができる。また、使用フラグUの値は1ビットだけであるがアクセス順序が古いか新しいかをしめすので、Uフラグ設定処理によって、スタートラインからエンドラインまでの各ラインについてアクセス順序を設定することができる。例えば、プロセッサ1は、キャッシュメモリ3に残しておきたいアドレスのデータをアドレス範囲として指定してU=1を設定するUフラグ設定コマンドを発行し、

以上説明してきたように、本実施の形態におけるキャッシュメモリによれば、Wフラグを設定することにより、W=1のキャッシュエントリーをキャッシュミス時には最古のキャ

ッシュエントリとして真っ先にリプレース対象とすることができる。また、使用フラグUの値は1ビットだけであるがアクセス順序が古いか新しいかをしめすので、Uフラグ設定処理によって、スタートラインからエンドラインまでの各ラインについてアクセス順序を設定することができる。例えば、プロセッサ1は、キャッシュメモリ3に残しておきたいアドレスのデータをアドレス範囲として指定してU=1を設定するUフラグ設定コマンドを発行すればよい。逆に、キャッシュメモリ3が追い出してもよいデータをアドレス範囲として指定してU=0を設定するUフラグ設定コマンドを発行すればよい。

[0103] <変形例>

(1) W=1のキャッシュエントリは真っ先にリプレース対象とされるが、制御部は、リプレースされるまでの間に、ダーティであればクリーニング(ライトバック)を行ってもよい。

(2) 図6(a)(b)(c)に示した各命令は、コンパイラによりプログラム中に挿入してもよい。その際、コンパイラは、例えば配列データの書き込みや、圧縮動画データをデコードする際のブロックデータの書き込み等、これ以上書き込みをしないプログラム位置に、上記各命令を挿入するようにすればよい。

産業上の利用可能性

[0104] 本発明は、メモリアクセスを高速化するためのキャッシュメモリに適しており、例えば、オンチップキャッシュメモリ、オフチップキャッシュメモリ、データキャッシュメモリ、命令キャッシュメモリ等に適している。

請求の範囲

- [1] キャッシュの単位データを保持するキャッシュエントリーに対応させて、当該キャッシュエントリーが有効であるか否かを示すバリッドフラグと、当該キャッシュエントリーに対する書き込みがなされたか否かを示すダーティフラグとを保持するフラグ保持手段と、
- プロセッサからの指示に基づき、前記バリッドフラグおよびダーティフラグの少なくとも一方をキャッシュエントリーの状態に反して改変する改変手段と
- を備えることを特徴とするキャッシュメモリ。
- [2] 前記改変手段は、メモリからデータをロードすることなく、キャッシュエントリーに対してタグとしてのアドレスを設定しバリッドフラグをセットすることを特徴とする請求項1記載のキャッシュメモリ。
- [3] 前記改変手段は、キャッシュエントリーに対してライトバックされていない書き換えられたデータを保持している状態でダーティフラグをリセットすることを特徴とする請求項2記載のキャッシュメモリ。
- [4] 前記キャッシュメモリは、さらに、
- プロセッサから指定されたアドレス範囲を保持する保持手段と、
- 保持されたアドレス範囲に属するデータを保持するキャッシュエントリーを特定する特定手段とを備え、
- 前記改変手段は、特定されたキャッシュエントリーに対して前記バリッドフラグおよびダーティフラグの少なくとも一方を改変することを特徴とする請求項2又は3記載のキャッシュメモリ。
- [5] 前記特定手段は、
- 前記アドレス範囲の先頭アドレスがラインデータの途中を指す場合、当該先頭アドレスを、前記アドレス範囲に含まれる先頭のラインを指すスタートラインアドレスに変換する第1変換手段と、
- 前記アドレス範囲の末尾アドレスがラインデータの途中を指す場合、当該末尾アドレスを、前記アドレス範囲に含まれる末尾のラインを指すエンドラインアドレスに変換する第2変換手段と、

前記スタートラインアドレスからエンドラインアドレスまでの各ラインアドレスに対応するデータを保持するキャッシュエントリーがあるか否かを判定する判定手段とを備えることを特徴とする請求項4記載のキャッシュメモリ。

[6] 前記改変手段は、

ダーティフラグのリセット指示付きメモリアクセス命令が実行されたことを検出する命令検出手段と、

当該命令によってアクセスがなされたキャッシュエントリーに対してダーティフラグをリセットするフラグ書き換え手段と

を備えることを特徴とする請求項1記載のキャッシュメモリ。

[7] 前記改変手段は、

バリッドフラグのリセット指示付きメモリアクセス命令が実行されたことを検出する命令検出手段と、

当該命令によってアクセスがなされたキャッシュエントリーに対してバリッドフラグをリセットするフラグ書き換え手段と

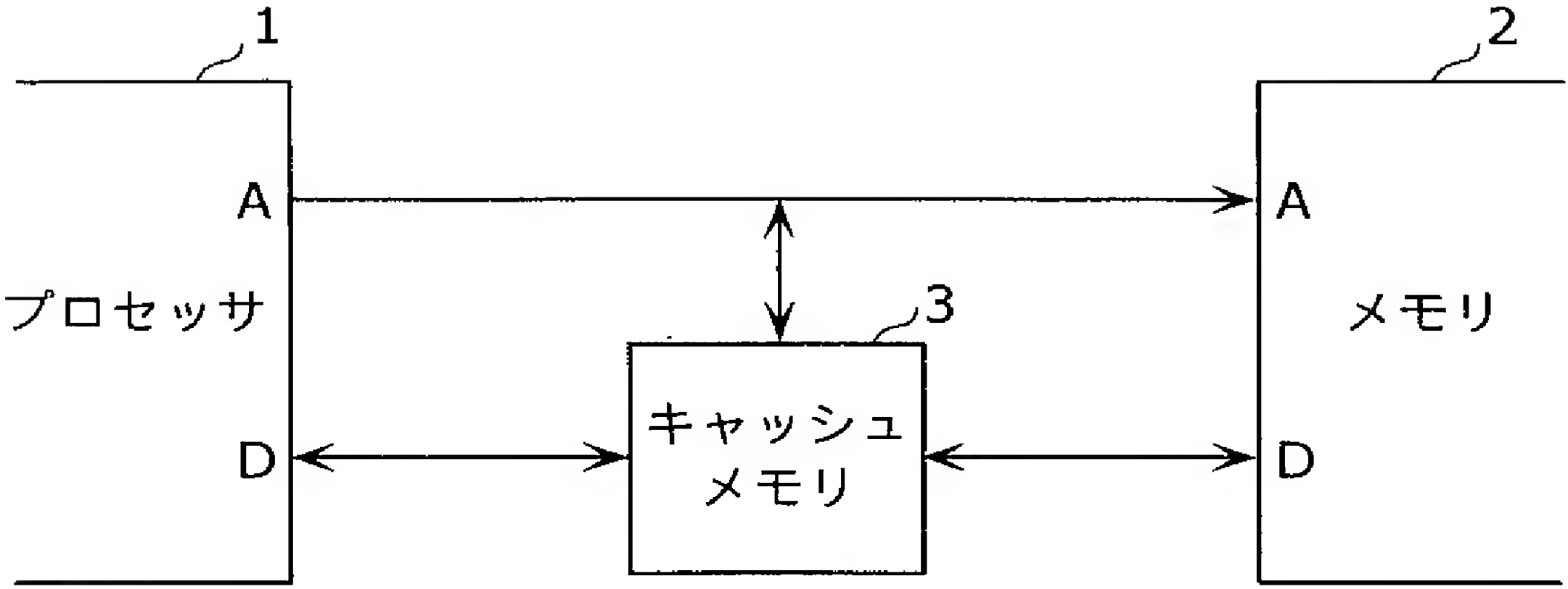
を備えることを特徴とする請求項1記載のキャッシュメモリ。

[8] キャッシュの単位データを保持するキャッシュエントリーに対応させて、当該キャッシュエントリーが有効であるか否かを示すバリッドフラグと、当該キャッシュエントリーに対する書き込みがなされたか否かを示すダーティフラグとを有するキャッシュメモリの制御方法であって、

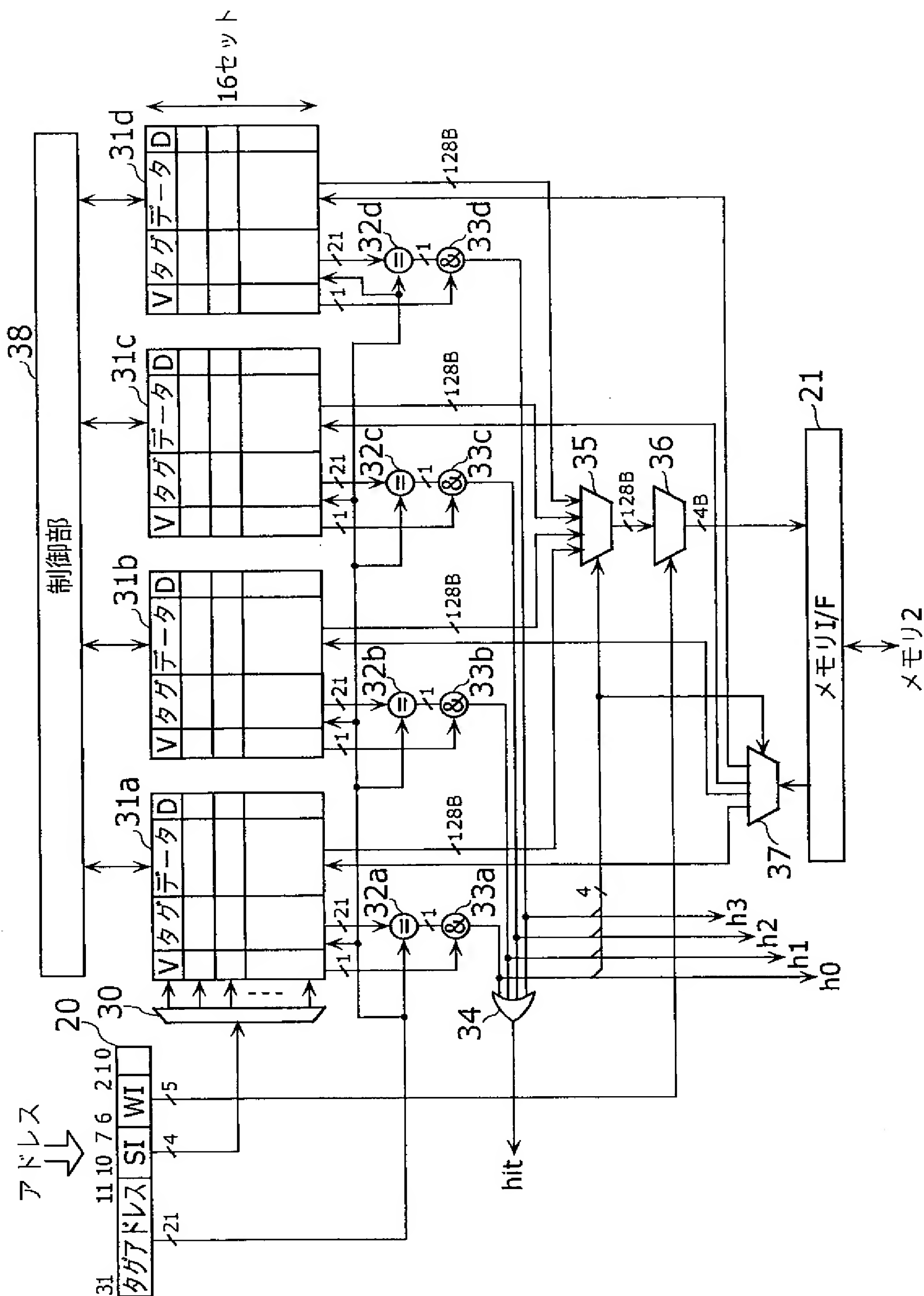
プロセッサの指示に基づき、メモリからデータをロードすることなく、キャッシュエントリーに対してタグとしてのアドレスを設定しバリッドフラグをセットするステップと、

プロセッサの指示に基づき、キャッシュエントリーに対してライトバックされていない書き換えられたデータを保持している状態でダーティフラグをリセットするステップとを有することを特徴とする制御方法。

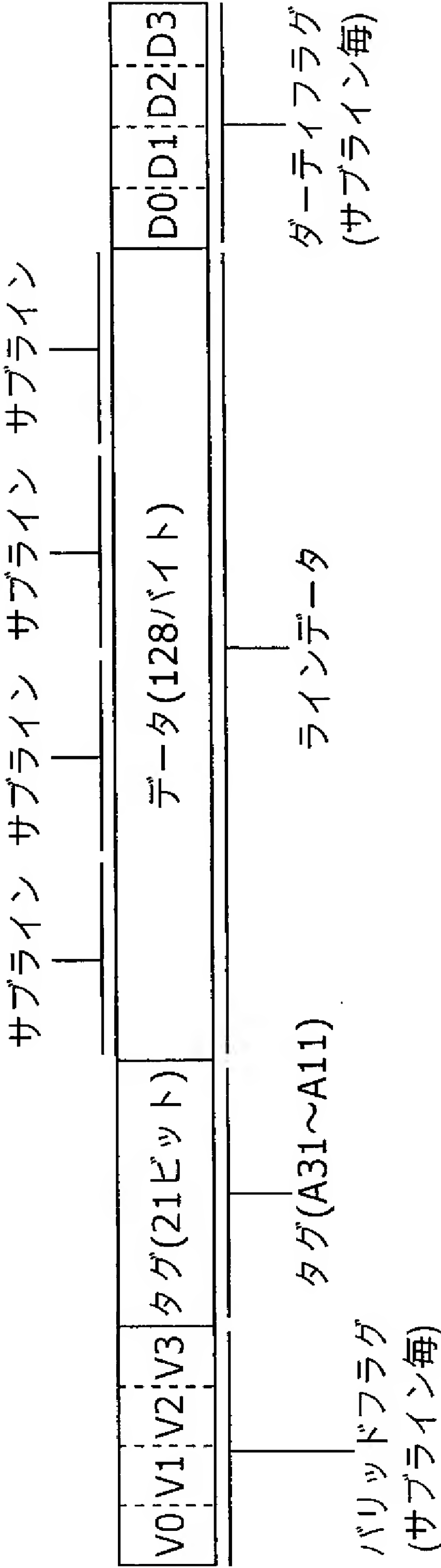
[図1]



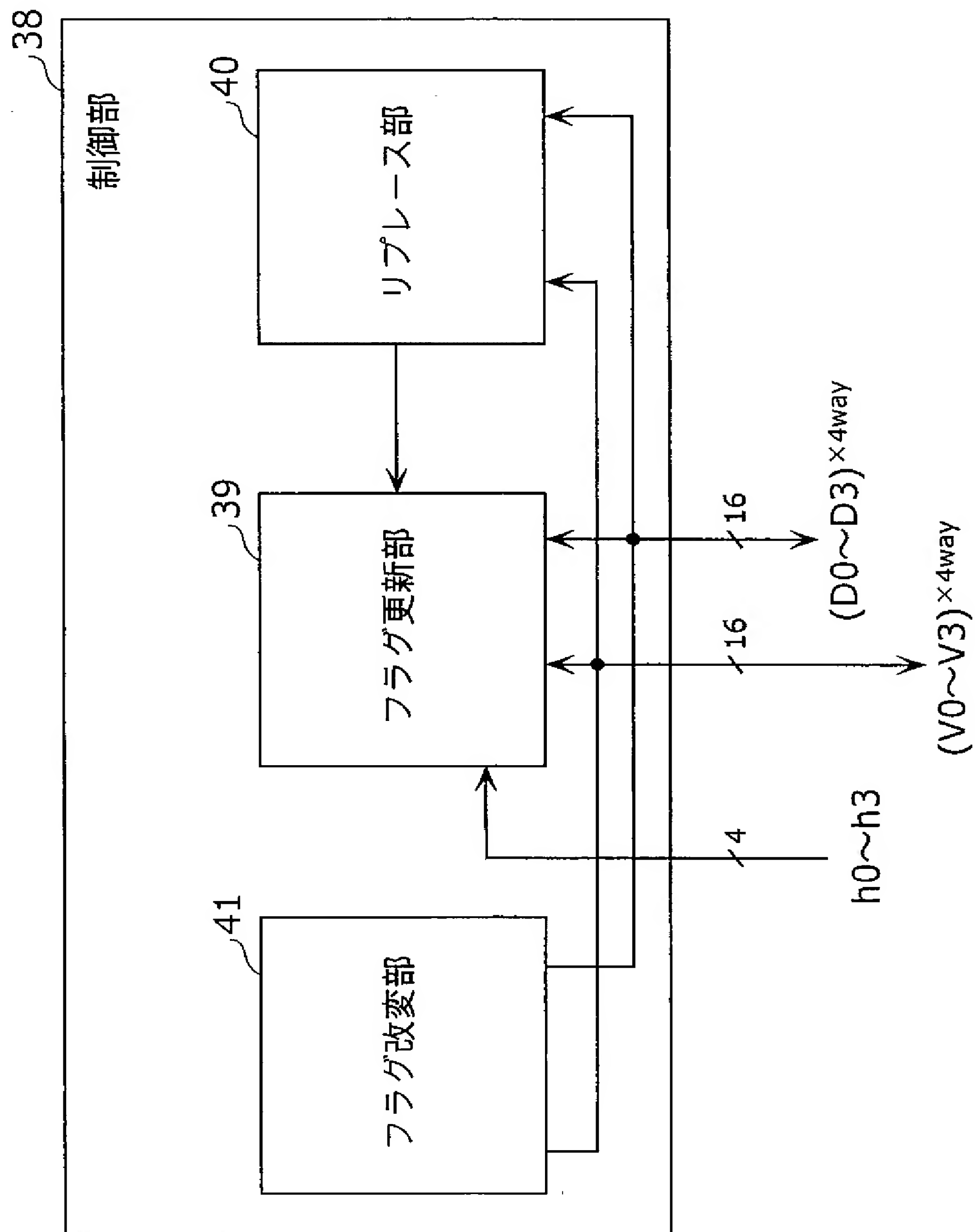
[図2]



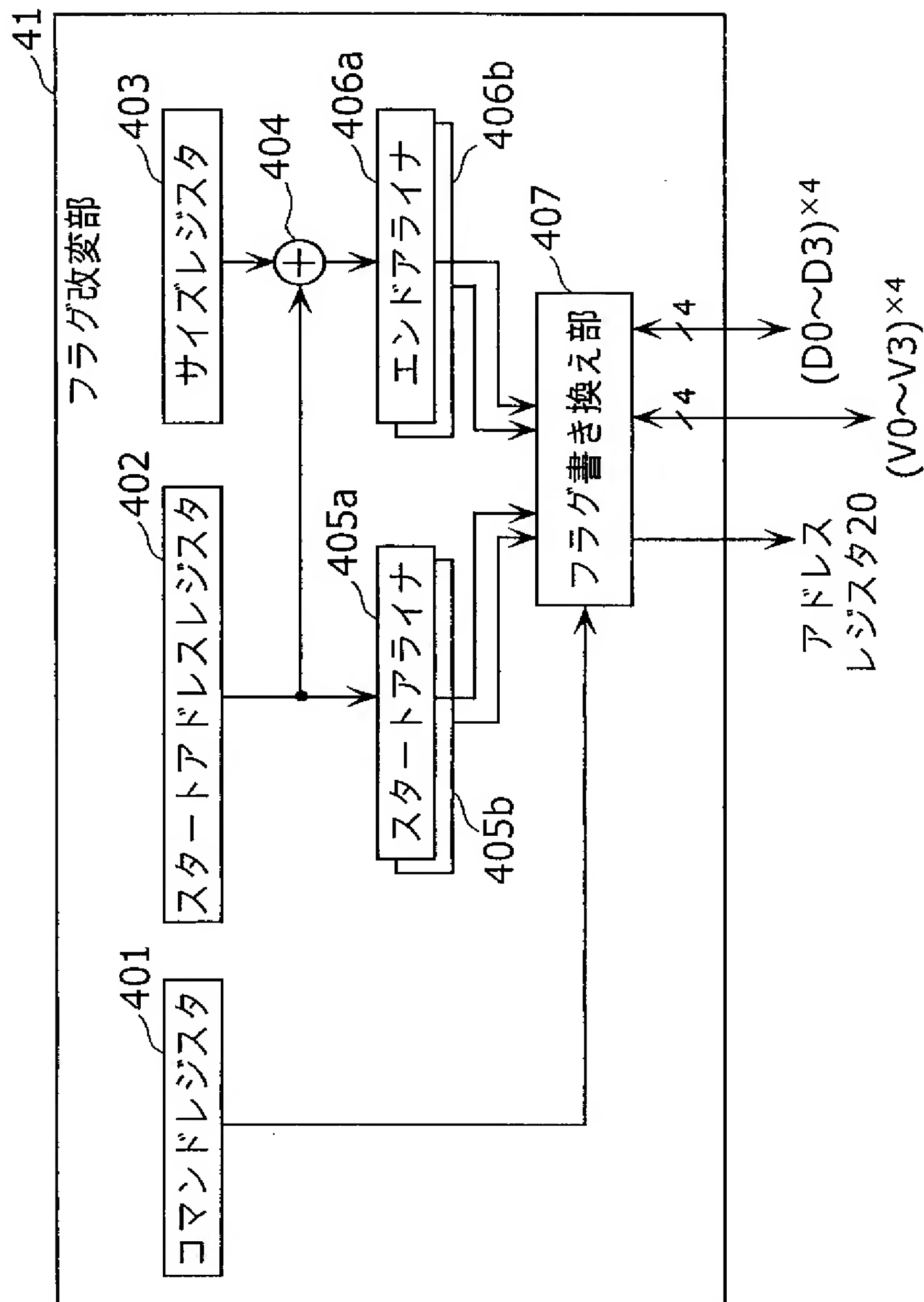
[図3]



[図4]



[図5]

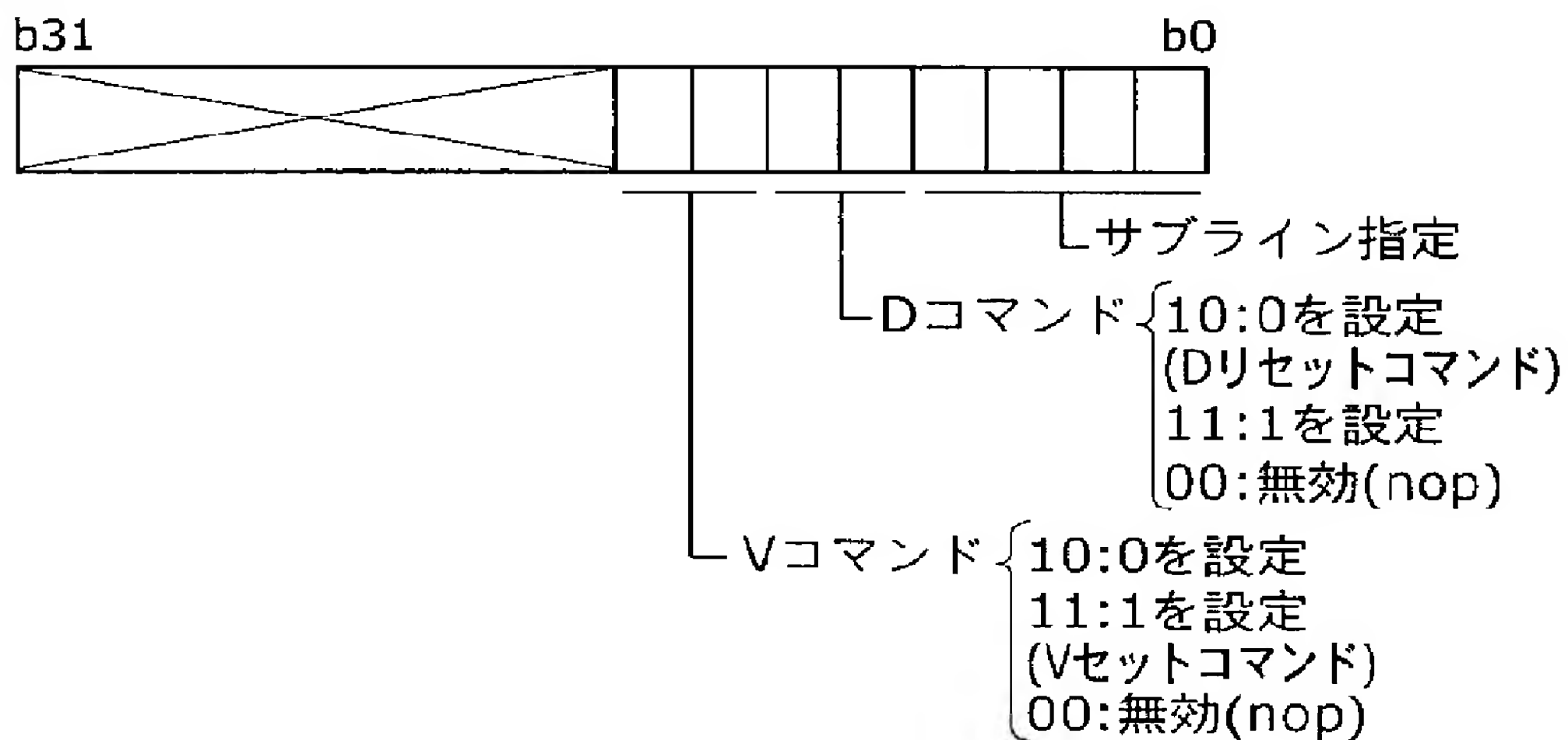


(a)

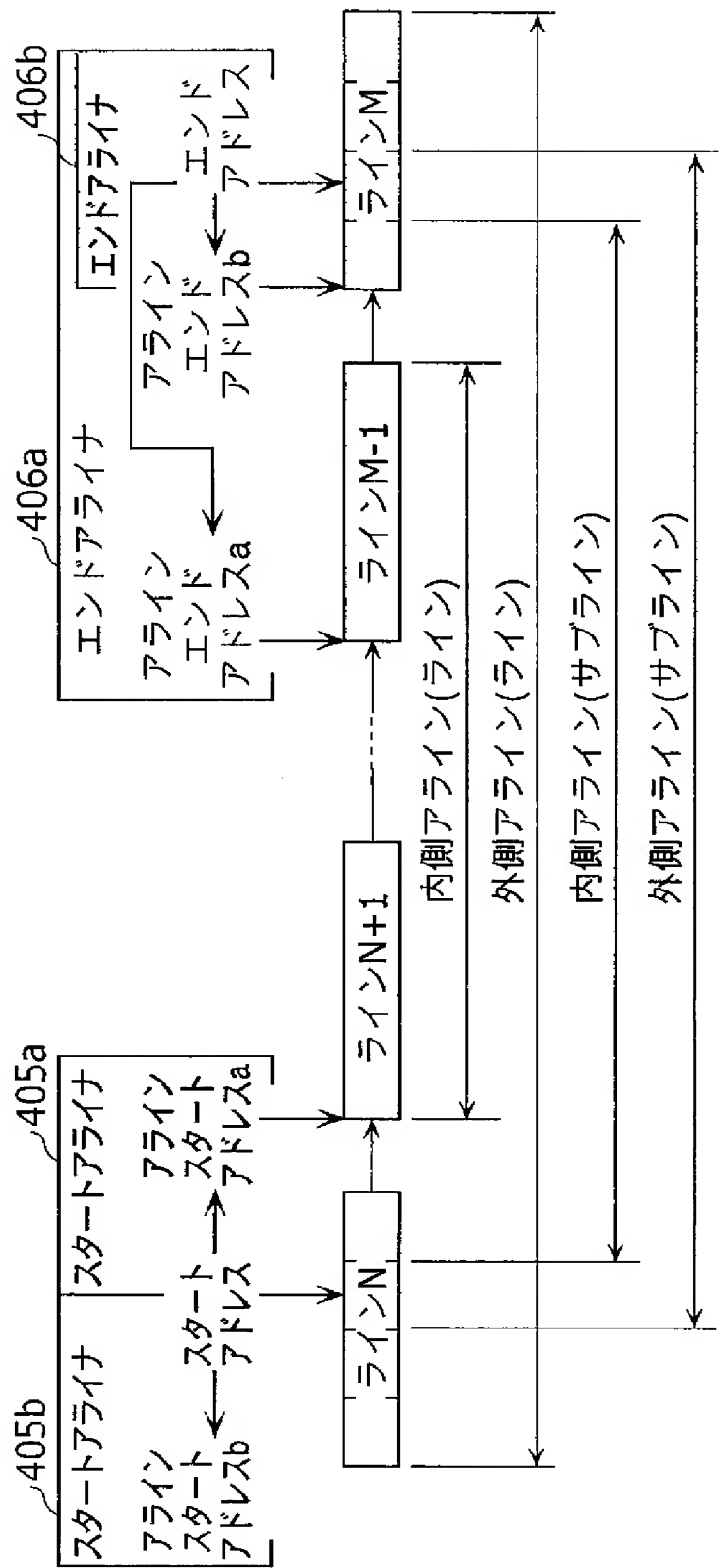
- └ (ソース)スタートアドレスを指定
- └ (デスティネーション)スタートアドレスレジスタを指定
- └ 転送命令

- └ (ソース)サイズを指定
- └ (デスティネーション)サイズレジスタを指定
- └ 転送命令

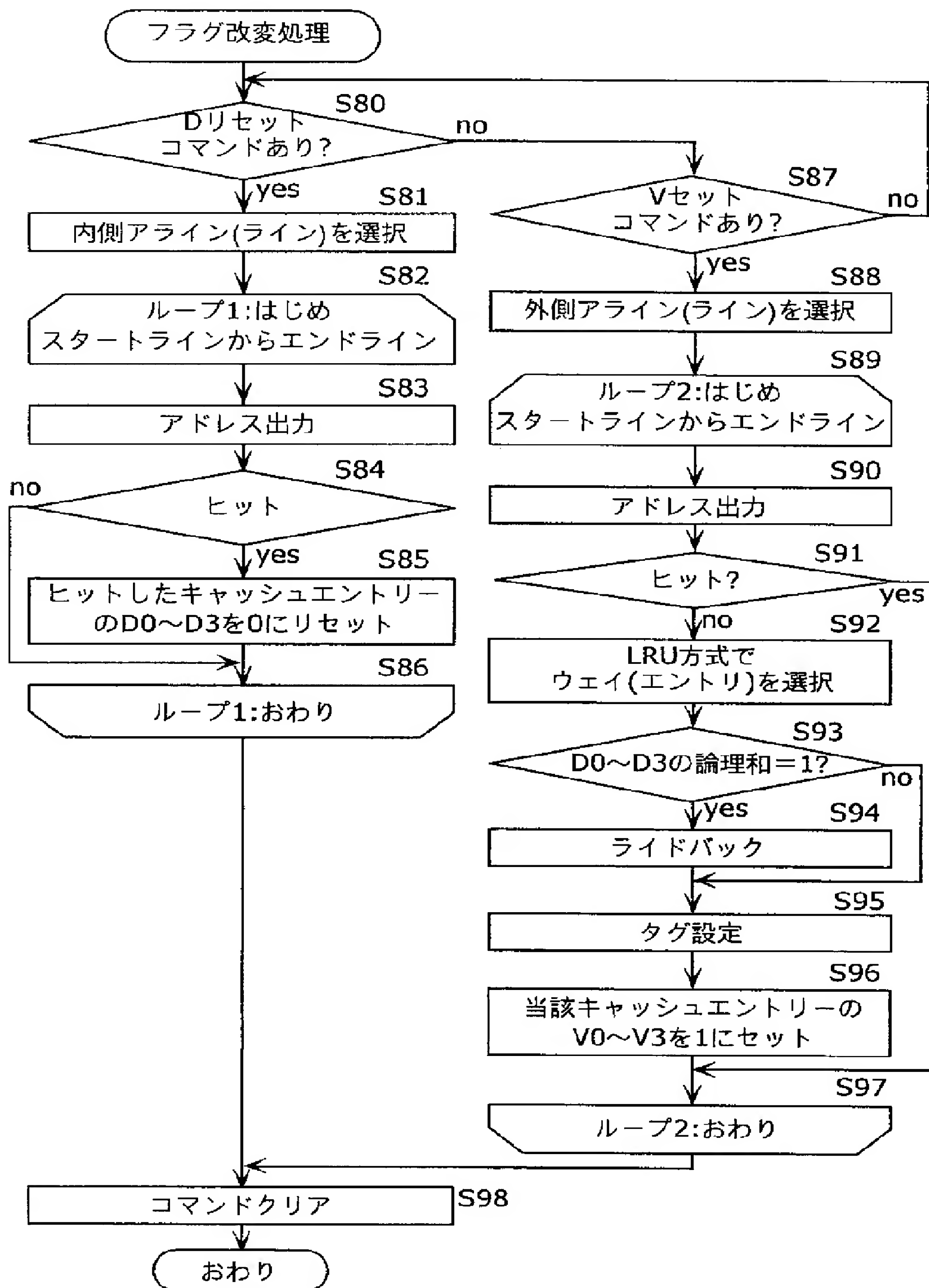
- └ (ソース) コマンド を指定
- └ (デスティネーション) コマンドレジスタを指定
- └ 転送命令



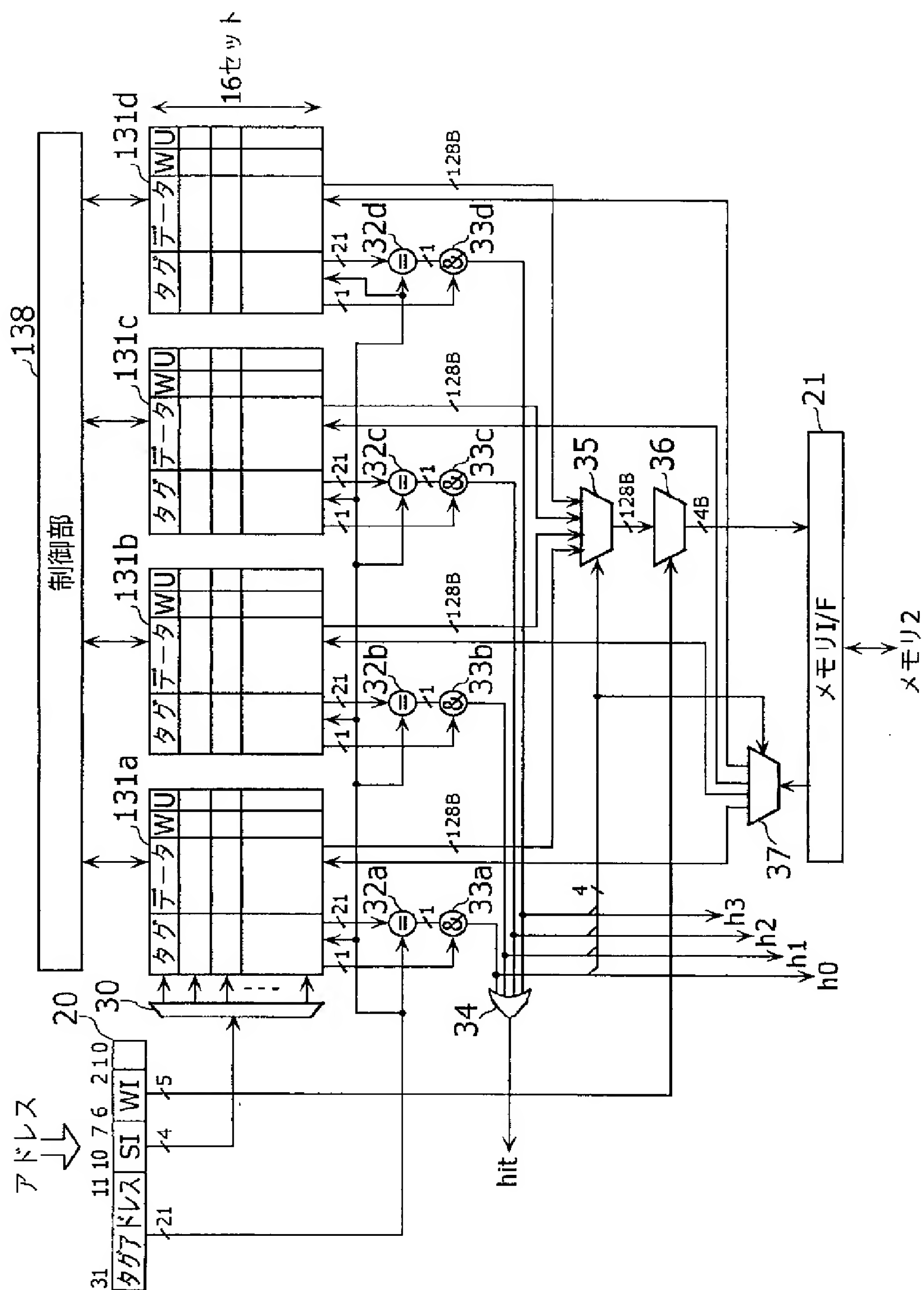
[図7]



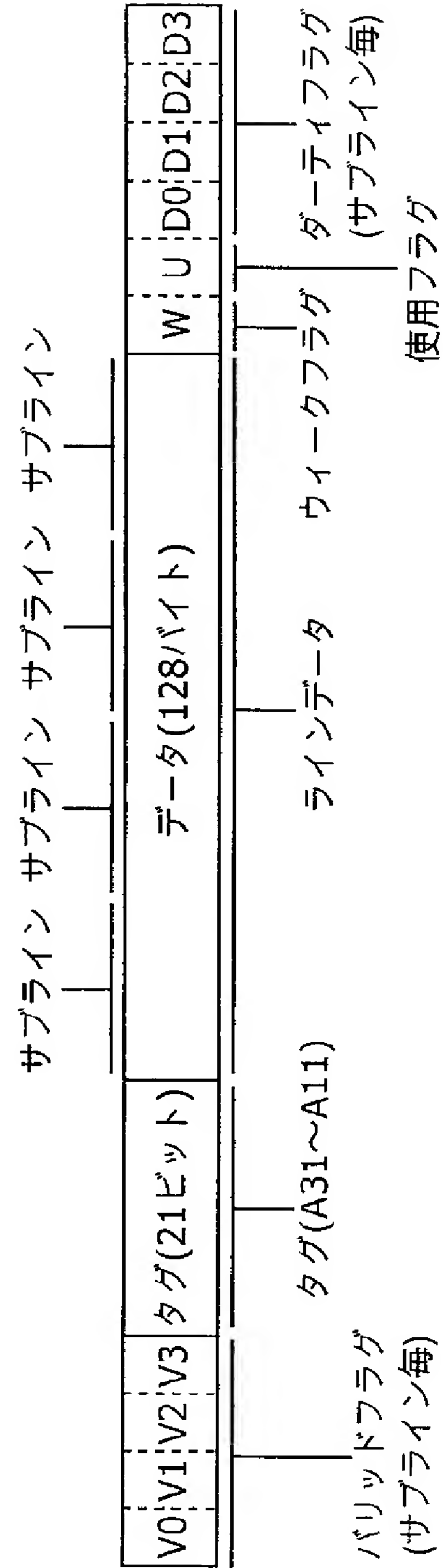
[図8]



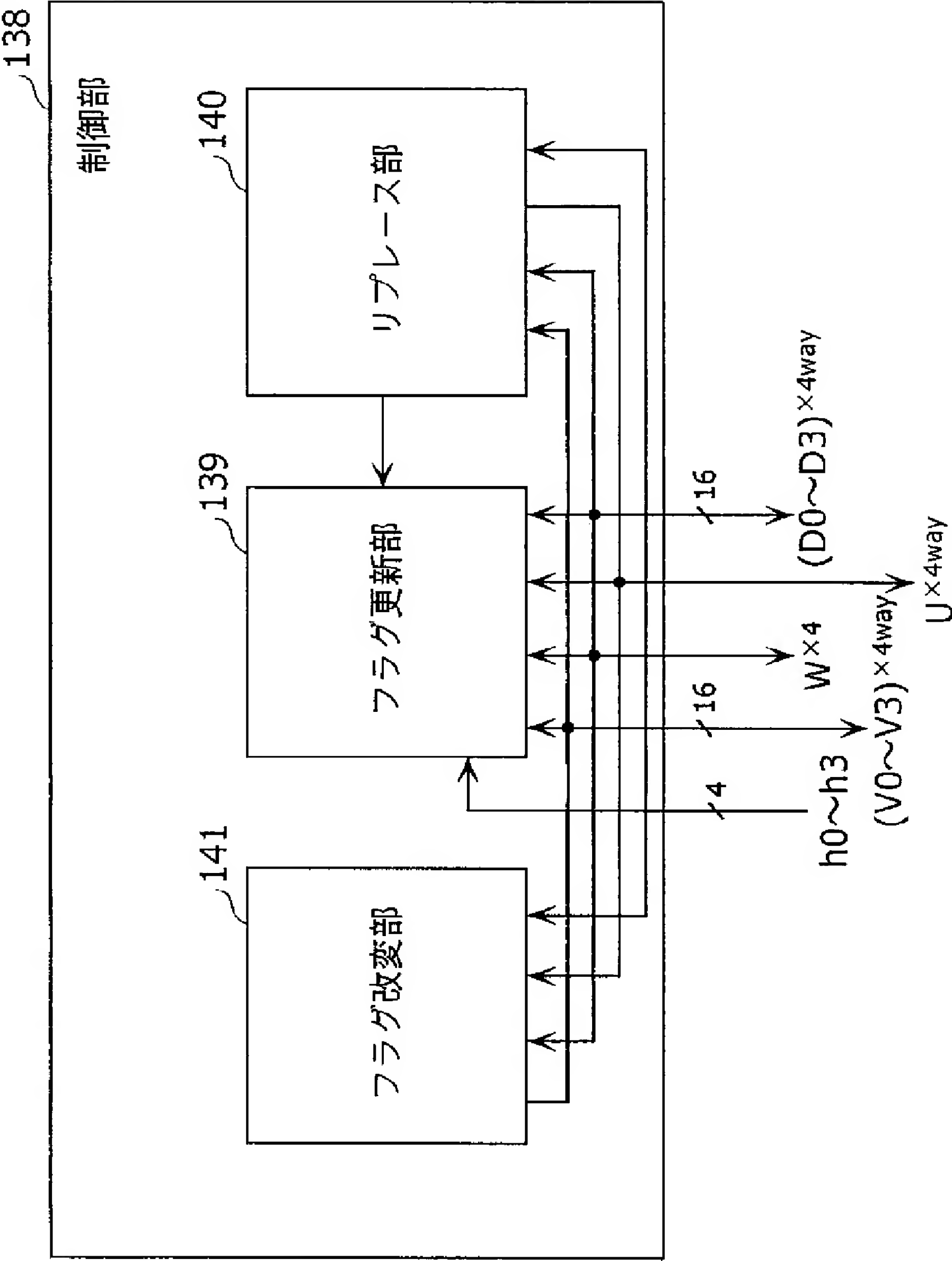
[図9]



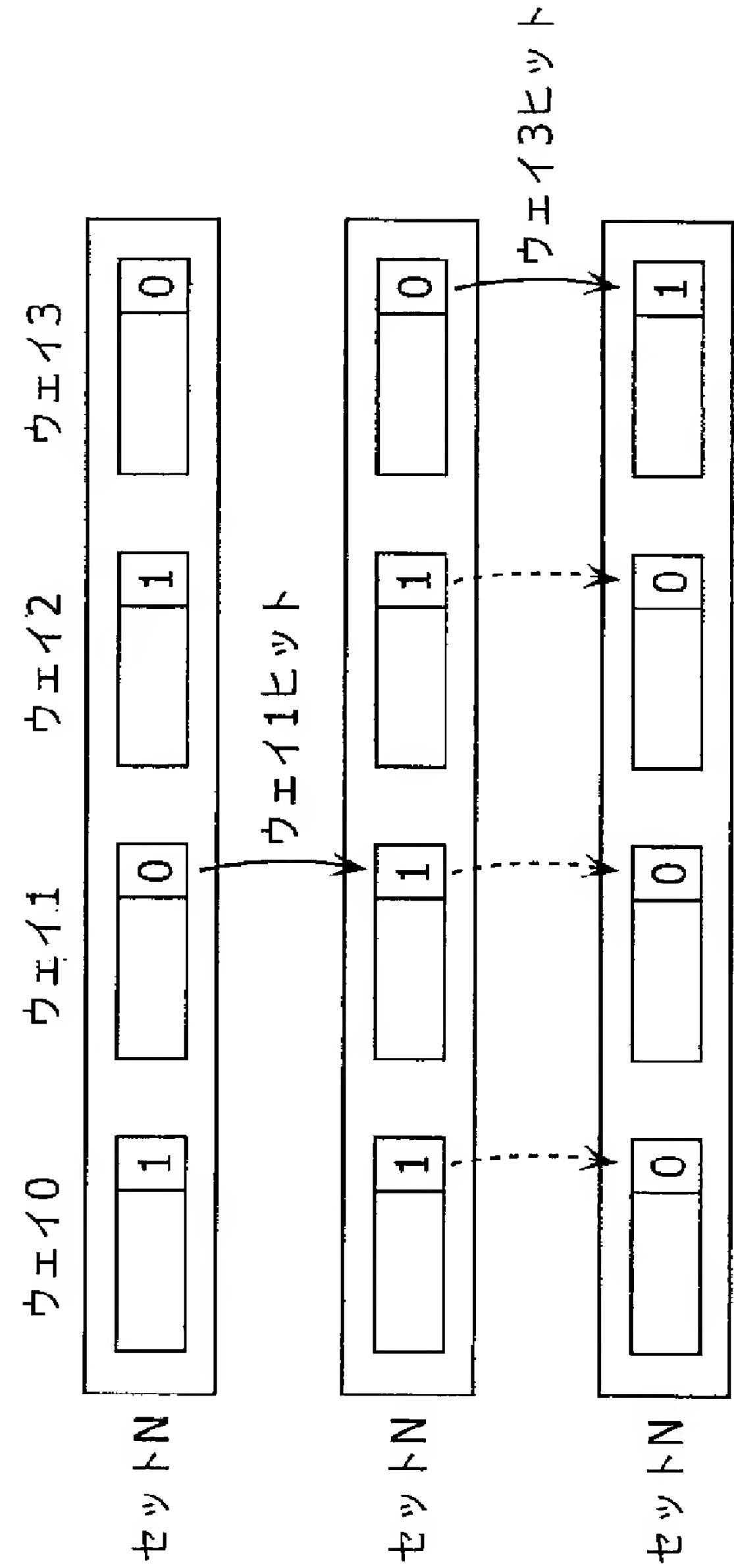
[図10]



[図11]

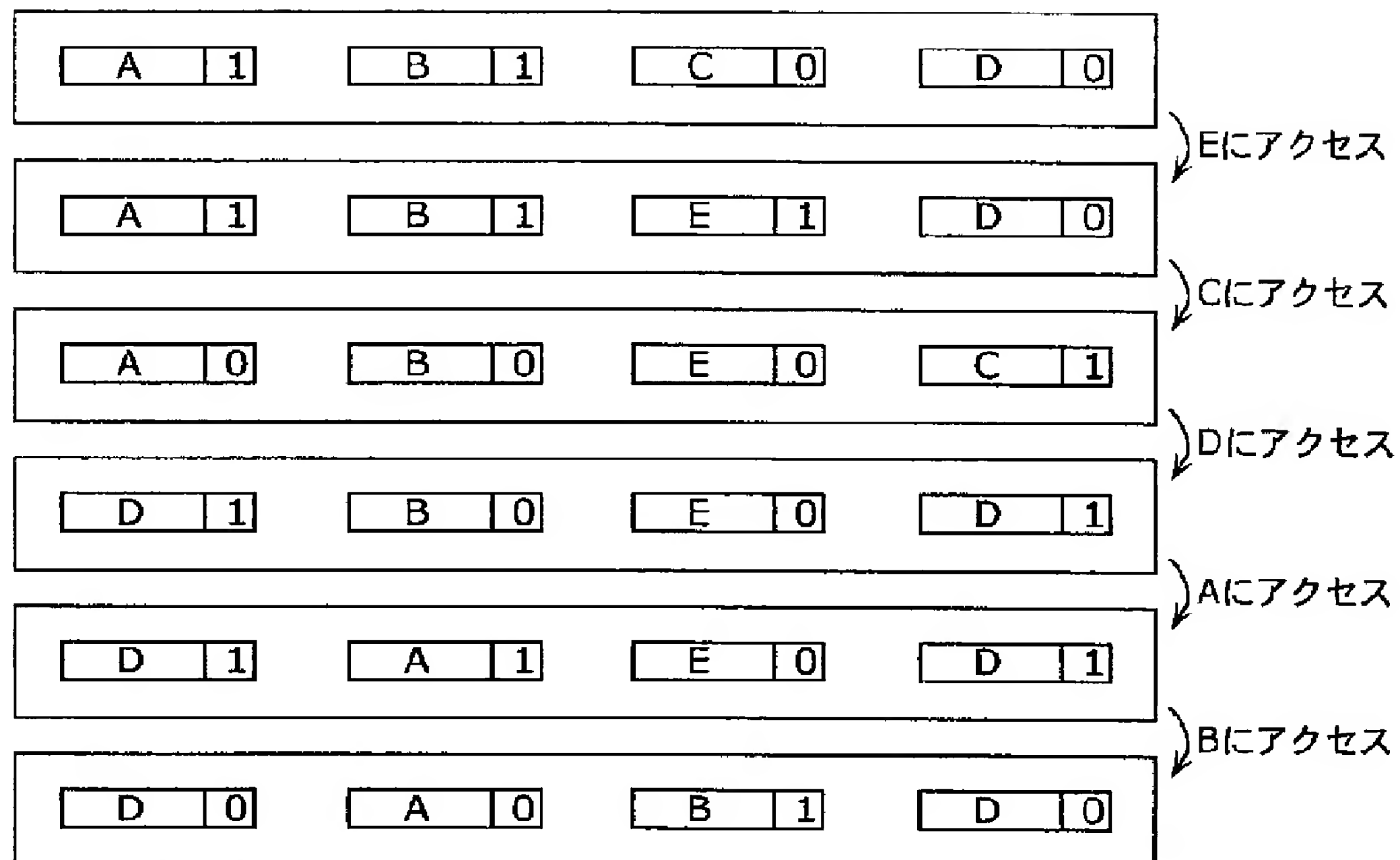


[図12]

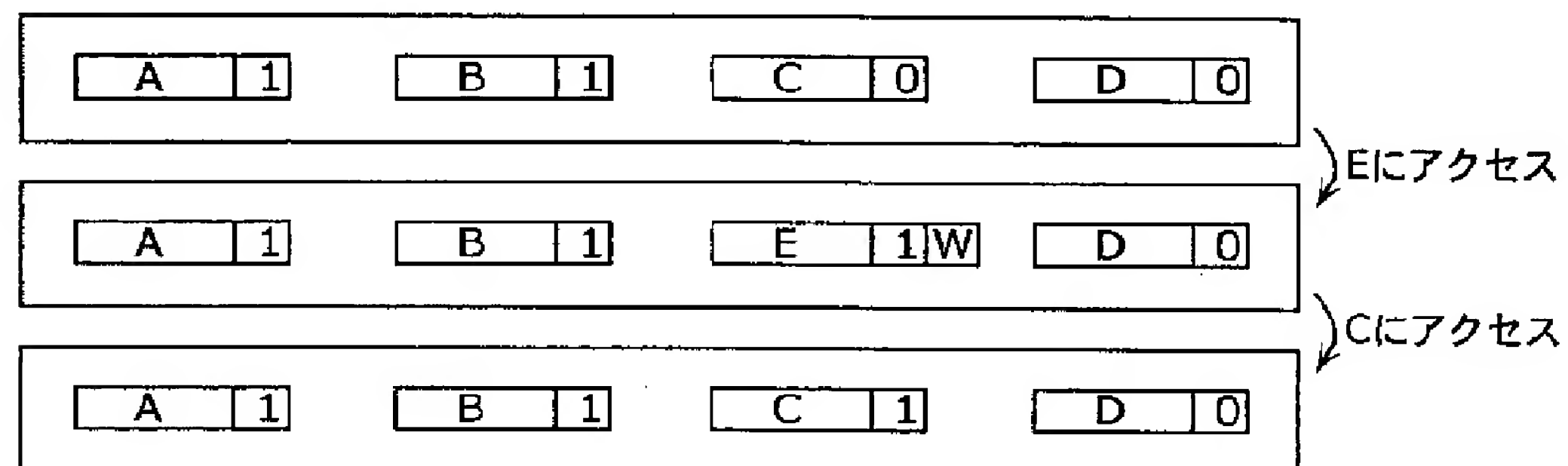


[図13]

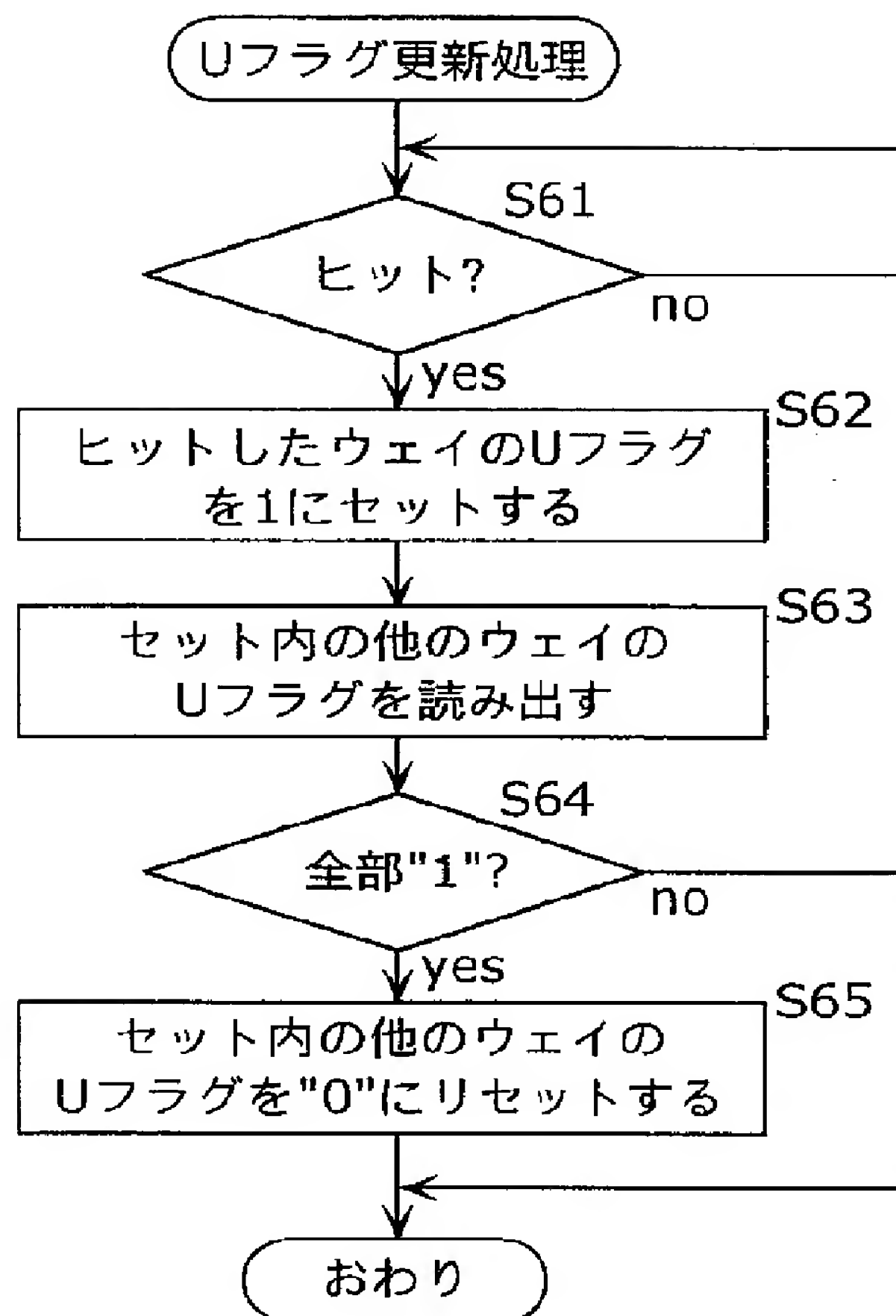
(a)



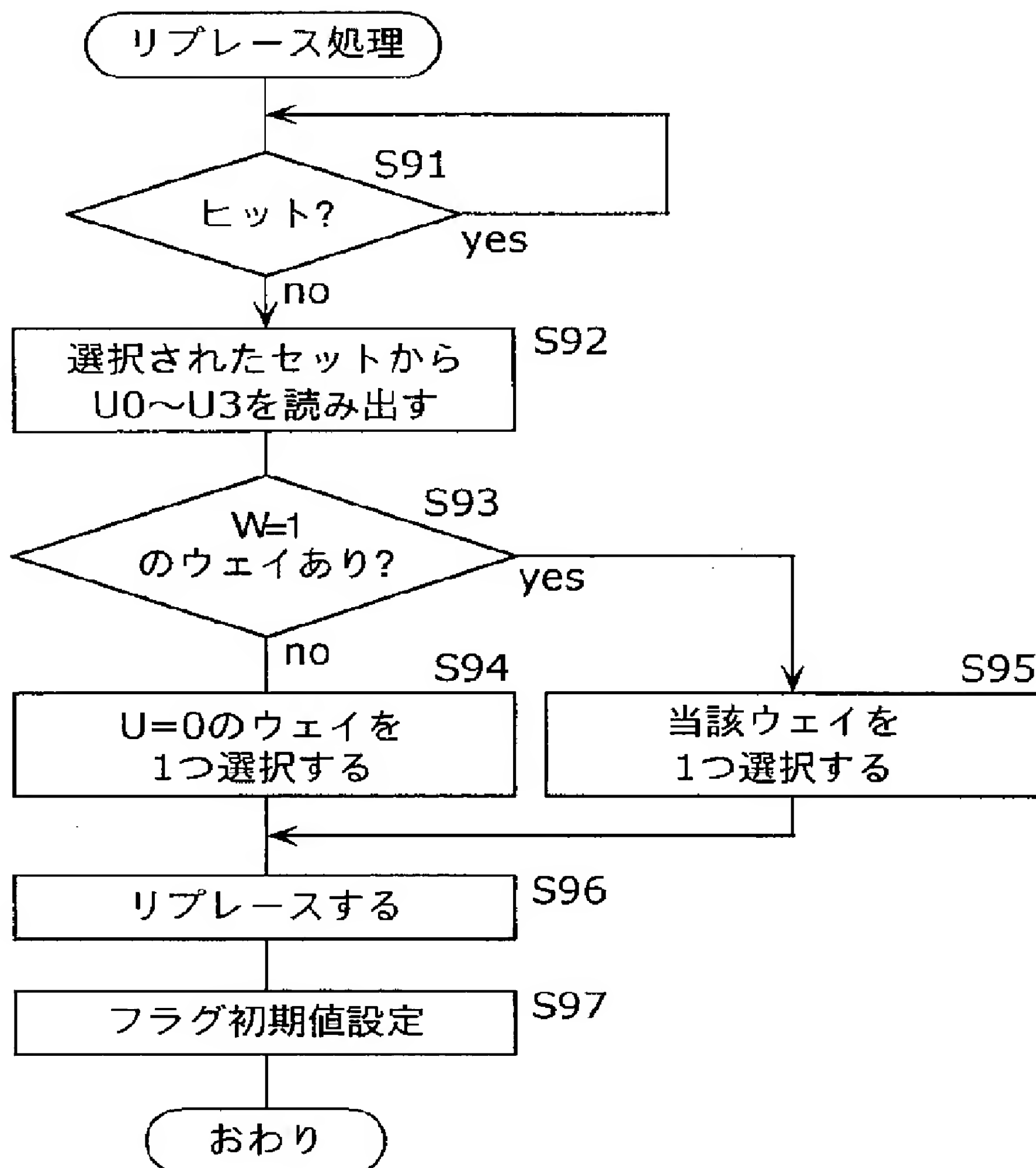
(b)



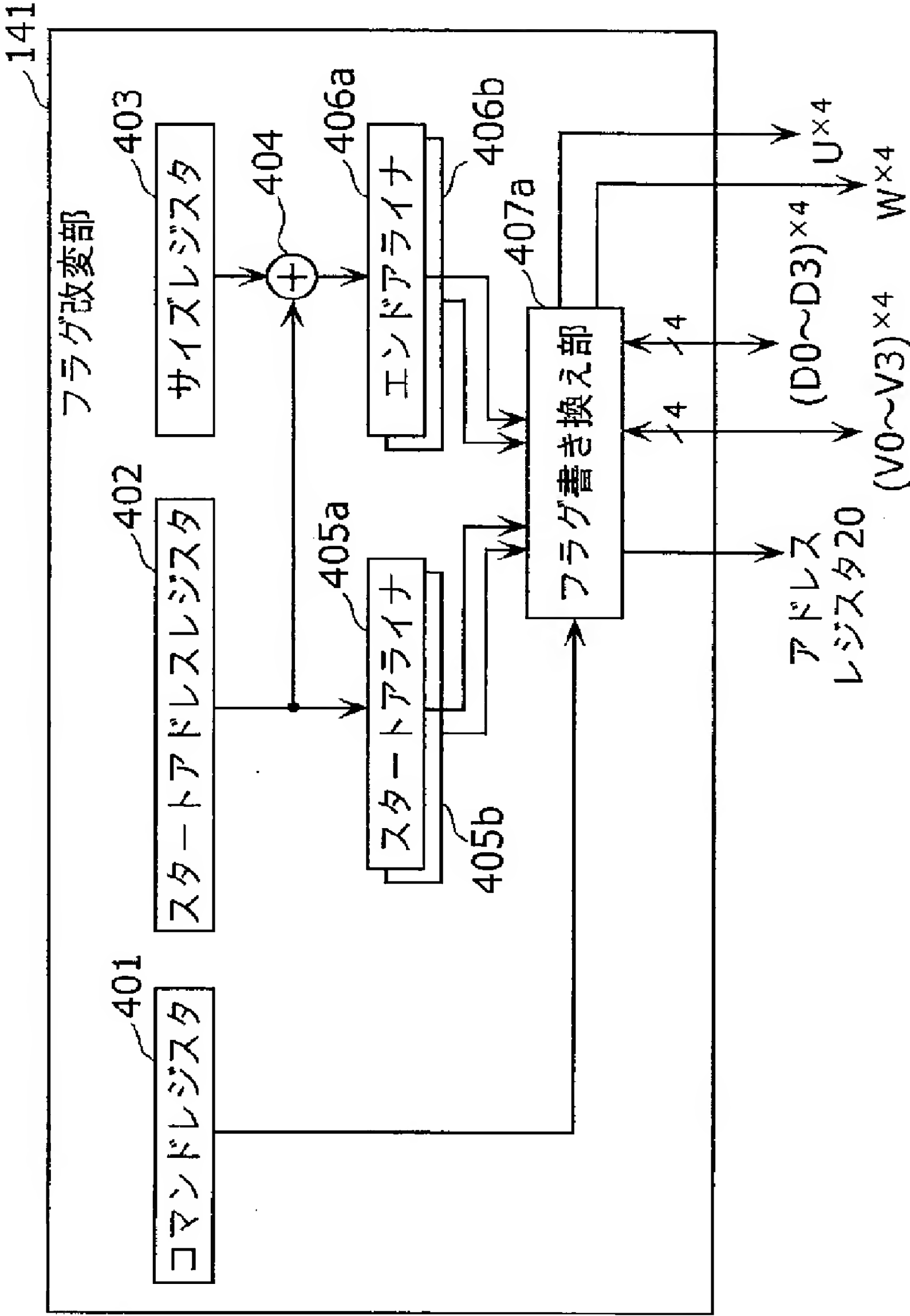
[図14]



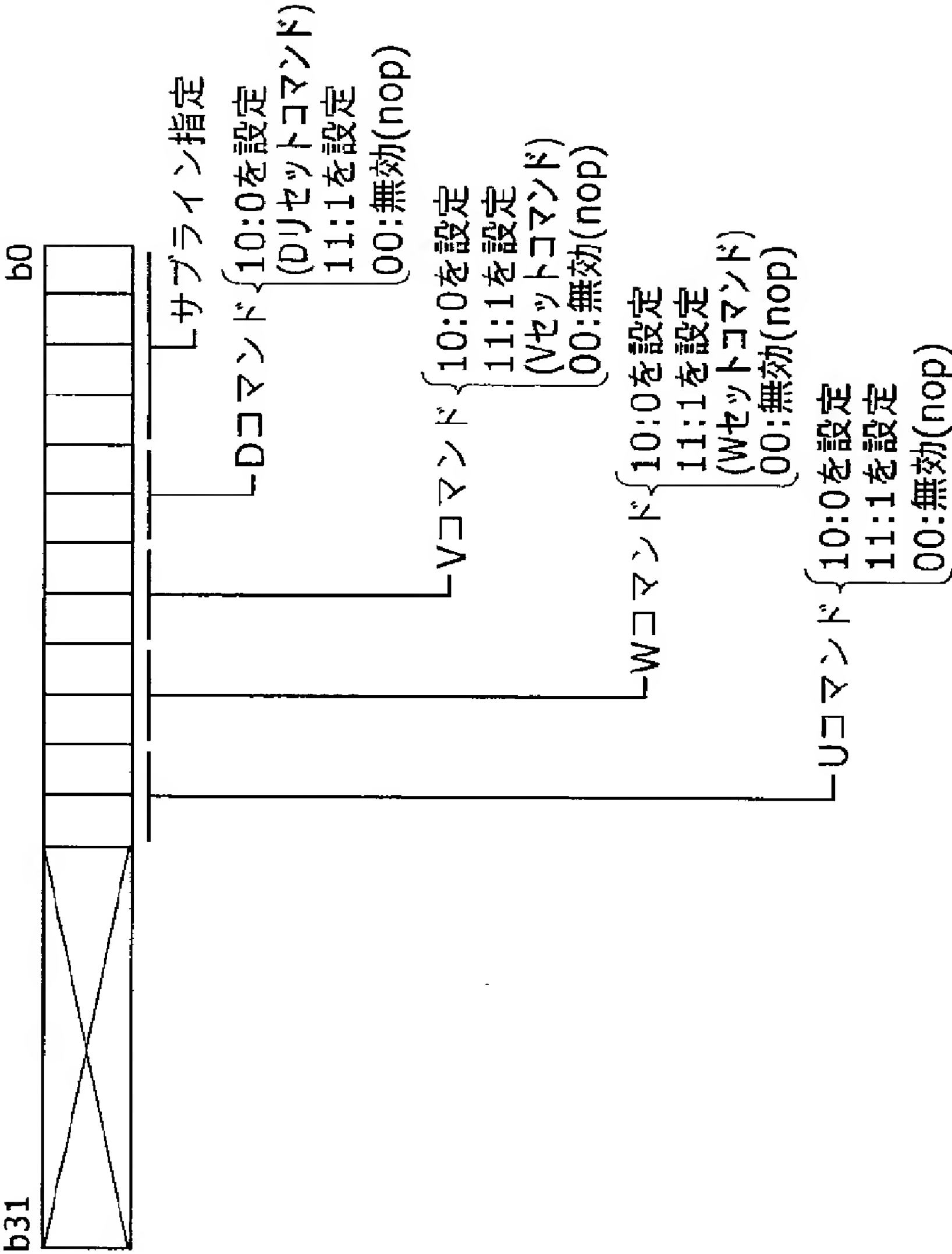
[図15]



[図16]



[図17]



```

graph TD
    Start([Wフラグ設定処理]) --> S81{コマンドあり}
    S81 -- yes --> S82[/ループ1:はじめ  
スタートラインからエンドライン/]
    S82 --> S83[アドレス出力]
    S83 --> S84{ヒット?}
    S84 -- yes --> S185[ヒットしたキャッシュエントリー  
にWフラグ設定する]
    S84 -- no --> S86
    S185 --> S86
    S86[/ループ1:おわり/] --> End([おわり])

```

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/019102

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl⁷ G06F12/08, G06F12/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl⁷ G06F12/08-12/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2005
Kokai Jitsuyo Shinan Koho 1971-2005 Toroku Jitsuyo Shinan Koho 1994-2005

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 2003-223360 A (Hitachi, Ltd.), 08 August, 2003 (08.08.03), Claims 1 to 7; Par. Nos. [0001] to [0023], [0067] to [0098]; Figs. 6 to 12 (Family: none)	1 3, 4, 5, 7, 8
X Y	JP 2000-29788 A (NEC Corp.), 28 January, 2000 (28.01.00), Par. Nos. [0060] to [0084]; Figs. 1, 2, 3, 6, 7 (Family: none)	1, 2 3, 4, 5, 8
X Y	JP 2001-222467 A (Matsushita Electric Industrial Co., Ltd.), 17 August, 2001 (17.08.01), (Family: none)	1, 2, 3, 8 4, 5

☒ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
04 April, 2005 (04.04.05)

Date of mailing of the international search report
26 April, 2005 (26.04.05)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2004/019102

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	JP 3-54649 A (Oki Electric Industry Co., Ltd.), 08 March, 1991 (08.03.91), Page 5, upper right column, line 4 to lower right column, line 5; Fig. 4(a) (Family: none)	1, 6 3, 4, 5, 7, 8
A	JP 7-84879 A (Toshiba Corp.), 31 March, 1995 (31.03.95), (Family: none)	1-8

A. 発明の属する分野の分類 (国際特許分類 (I P C))

Int.Cl.⁷ G06F12/08
G06F12/12

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (I P C))

Int.Cl.⁷ G06F12/08-12/12

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2005年
日本国実用新案登録公報	1996-2005年
日本国登録実用新案公報	1994-2005年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X Y	J P 2003-223360 A (株式会社日立製作所) 2003.08.08, 【請求項1】-【請求項7】, 【0001】-【0023】、【0067】-【0098】, 【図6】-【図12】(ファミリーなし)	1 3, 4, 5, 7, 8

☒ C欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

04.04.2005

国際調査報告の発送日

26.04.2005

国際調査機関の名称及びあて先

日本国特許庁 (I S A/J P)

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

清木 泰

電話番号 03-3581-1101 内線 3586

5 N

9643

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
X Y	J P 2000-29788 A (日本電気株式会社) 2000.01.28, 【0060】-【0084】, 【図1】、【図2】、【図3】、【図6】、【図7】 (ファミリーなし)	1, 2 3, 4, 5, 8
X Y	J P 2001-222467 A (松下電器産業株式会社) 2001.08.17 (ファミリーなし)	1, 2, 3, 8 4, 5
X Y	J P 3-54649 A (沖電気株式会社) 1991.03.08, 第5頁右上欄第4行-右下欄第5行, 第4図 (a) (ファミリーなし)	1, 6 3, 4, 5, 7, 8
A	J P 7-84879 A (株式会社東芝) 1995.03.31 (ファミリーなし)	1-8